

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:04 ; Search time 162.447 Seconds  
(without alignments)  
359.949 Million cell updates/sec

Title: US-10-057-813-14

Perfect score: 861

Sequence: 1 MEHYKAGSVELPAPSPMPQ.....SMPSSCGPRRRARDTRTS 163

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_23Sep04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1980s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	861	100.0	163	4	AAG89339	Aag89339 Human sec
2	861	100.0	163	6	ABU00014	Abu00014 Human nov
3	861	100.0	219	8	ADM90988	Adm90988 Human pha
4	813	94.4	163	2	AA336104	Aay36104 Extended
5	813	94.4	163	8	ADP19412	Adp19412 Human sec
6	402	46.7	79	2	AA111950	Aay11950 Human 5'
7	395	45.9	78	3	AAG03599	Amg03599 Human sec
8	339	39.4	211	4	AA338973	Am38973 Human pol
9	232.5	27.0	205	4	ABB64410	Abb64410 Drosophil
10	206	23.9	350	3	AAG16838	Aagi6838 Arabidops
11	206	23.9	350	3	AAG1875	Aags1875 Arabidops
12	206	23.9	350	3	AAG51847	Aags1847 Arabidops
13	206	23.9	350	8	ADN72883	Adn72883 Thale cre
14	206	23.9	350	8	ADN72183	Adn72183 Thale cre
15	206	23.9	356	3	AAG51846	Aags1846 Arabidops
16	198	23.0	267	3	AAG33136	Aag33136 Zea mays
17	192	22.3	259	3	AAG34072	Aag34072 Zea mays
18	192	22.3	313	3	AAG34071	Aag34071 Zea mays
19	186	21.6	239	3	AAG33137	Aag33137 Zea mays
20	171.5	19.9	314	3	AAG51848	Aags1848 Arabidops
21	171.5	19.9	314	3	AAG16839	Aagi6839 Arabidops
22	171.5	19.9	314	3	AAG51876	Aags1876 Arabidops
23	155.5	18.1	207	3	AAG34073	Aag34073 Zea mays
24	155.5	18.1	298	3	AAG51877	Aags1877 Arabidops
25	155.5	18.1	298	3	AAG16840	Aagi6840 Arabidops

26	154	17.9	215	3	AAG33138	Aag33138 Zea mays
27	150.5	17.5	99	3	AAG36977	Aag36977 Arabidops
28	135	15.7	74	3	AAG36978	Aag36978 Arabidops
29	104	12.1	124	8	ADP29604	Adp29604 Human sec
30	104	12.1	124	8	ADP29613	Adp29613 Human sec
31	104	12.1	124	8	ADP29615	Adp29615 Human sec
32	104	12.1	124	8	ADP29610	Adp29610 Human sec
33	104	12.1	124	8	ADP29607	Adp29607 Human sec
34	91.5	10.6	791	4	ABG23210	Abg23210 Novel hum
35	91.5	10.6	950	7	ABO77181	AbO77181 Pseudomon
36	91	10.6	345	7	ABO70060	AbO70060 Pseudomon
37	91	10.6	321	7	ABO68758	AbO68758 Pseudomon
38	90.5	10.5	330	7	ABO70617	AbO70617 Pseudomon
39	90	10.5	429	7	ABO71617	AbO71617 Pseudomon
40	89.5	10.4	432	7	ABO74876	AbO74876 Pseudomon
41	88	10.2	393	4	AAU27871	Aau27871 Fruit fly
42	87	10.1	374	7	ABO80084	AbO80084 Pseudomon
43	86.5	10.0	174	6	ADA55419	Ada55419 Human pro
44	86.5	10.0	359	4	AAM25759	Aam25759 Human pro
45	86.5	10.0	371	4	ABG05200	Abg05200 Novel hum

#### ALIGNMENTS

RESULT 1  
AAG89339  
ID AAG89339 standard; protein; 163 AA.  
XX AC AAG89339;  
XX DT 11-SRP-2001 (first entry)  
XX DE Human secreted protein, SEQ ID NO: 459.  
XX KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
KW GENSET.  
XX OS Homo sapiens.  
XX PN WO200142451-A2.  
XX PD 14-JUN-2001.  
XX PF 07-DEC-2000; 2000WO-IB001938.  
XX PR 08-DEC-1999; 99US-0169629P.  
PR 06-MAR-2000; 2000US-0187470P.  
(GEST ) GENSET.  
Dumas Milne Edwards J, Bougueleret L, Jobert S;  
WPI; 2001-367870/38.  
N-PSDB; AAH64942.  
Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.  
Claim 21; Page 910-911; 921pp; English.  
The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar nucleic acid sequences in samples, and hence to determine which  
 CC patients may be in need of restorative therapy. The GENSER polypeptides  
 CC may be used as antigens in the production of antibodies and in assays to  
 CC identify modulators (agonists and antagonists) of GENSER polypeptide  
 CC expression and activity. The present sequence is a GENSER polypeptide of  
 CC the invention

XX SQ Sequence 163 AA;

Query Match 100.0%; Score 861; DB 4; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-82;  
 Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSSARHVVFSGS 60  
 DB 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSSARHVVFSGS 60  
 QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120  
 DB 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120  
 QY 121 LSRDPLDPNECGYQPPGAPPGGLGMPSSSCGPRRRRRARDTRS 163  
 DB 121 LSRDPLDPNECGYQPPGAPPGGLGMPSSSCGPRRRRRARDTRS 163

# RESULT 2

ABU00014  
 ID ABU00014 standard; protein; 163 AA.

AC ABU00014;

DT 17-JAN-2003 (first entry)

DE Human novel polypeptide #107.

XX Human; genetic disorder; gene mapping; medical imaging; cancer;  
 KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;  
 KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;  
 KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;  
 KW fungal infection; bacterial infection; autoimmune disease; diabetes;  
 KW atopic dermatitis.

OS Homo sapiens.

XX WO200274961-A1.

XX 26-SEP-2002.

XX 14-MAR-2002; 2002WO-US005109.

XX 15-MAR-2001; 2001US-00810173.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich R, Auendi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2003-040556/03.

XX N-PSDB; ABX05092.

XX New isolated polypeptides and polynucleotides, useful for preventing,  
 PT treating or ameliorating medical conditions, such as cancer,  
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative  
 PT disorders, and infections.

XX Claim 9; SEQ ID NO 633; 235pp; English.

XX The invention relates to human polynucleotides and the polypeptides they  
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,  
 CC forensics, gene mapping, medical imaging, identification of mutations

CC responsible for genetic disorders or other traits, assessing biodiversity  
 CC and producing many other types of data and products dependent on DNA and  
 CC amino acid sequences. They are also useful for preventing, treating or  
 CC ameliorating medical conditions, such as cancer, neurodegenerative  
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell  
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,  
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or  
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).  
 CC Sequences ABG99888-ABG9989 and ABU00010-ABU00433 represent human  
 CC polypeptides of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied by the European Patent Office

XX SQ Sequence 163 AA;

Query Match 100.0%; Score 861; DB 6; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-82;  
 Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSSARHVVFSGS 60

DB 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSSARHVVFSGS 60

QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120

DB 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120

QY 121 LSRDPLDPNECGYQPPGAPPGGLGMPSSSCGPRRRRRARDTRS 163

DB 121 LSRDPLDPNECGYQPPGAPPGGLGMPSSSCGPRRRRRARDTRS 163

# RESULT 3

ADM90988

ID ADM90988 standard; protein; 219 AA.

XX ADM90988;

XX 03-JUN-2004 (first entry)

DE Human pharmaceutically useful protein SeqID 381.

XX human; cancer; haematopoiesis; thrombosis; anaemia;  
 KW cardiovascular disorder; ischaemic heart disease;  
 KW acute myocardial infarction; respiratory disease; asthma; pneumonia;  
 KW cystic fibrosis; chronic renal failure; glomerulopathy;  
 KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;  
 KW HIV infection; systemic lupus erythematosus; endocrine system;  
 KW diabetes mellitus; epilepsy; Alzheimer's disease;  
 KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;  
 KW fungal; parasitic; viral infection; cytostatic; anticoagulant;  
 KW thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;  
 KW antiinflammatory; nephrotropic; dermatologic; antitumor; hepatotropic;  
 KW immunosuppressive; antiallergic; dermatological; antirheumatic;  
 KW antiarthritic; antidiabetic; anticonvulsant; neuroprotective; nootropic;  
 KW antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;  
 XX gene therapy; vaccine.

OS Homo sapiens.

XX WO2004020595-A2.

XX 11-MAR-2004.

XX 28-AUG-2003; 2003WO-US027107.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410951P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411111P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.  
PA (DNAP-) DNAFORM KK.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K;  
XX  
XX WPI; 2004-257410/24.  
XX  
DR N-PSDB; ADM90779, ADM91197.  
XX  
XX  
PT New human polynucleotides and polypeptides, useful for diagnosing,  
PT preventing and treating proliferative disorders, immune disorders,  
PT cardiovascular disorders, or bacterial, fungal, parasitic and viral  
PT diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 381; 254pp; English.  
XX  
XX This invention relates to novel isolated human polynucleotides and the  
XX encoded proteins thereof. Specifically, it refers to proteases, kinases,  
XX phosphatases, secreted and transmembrane proteins, as well as the derived  
XX peptide fragments, which can be used to develop antibodies and screen for  
XX small molecule agonists and antagonists that can modulate their  
XX activities. The present invention describes polypeptides,  
XX polynucleotides, vectors and host cells useful for diagnosing, preventing  
XX and treating proliferative disorders, e.g. cancer, disorders of  
XX haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,  
XX e.g. ischaemic heart disease and acute myocardial infarction, respiratory  
XX diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the  
XX kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,  
XX gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune  
XX disorders, e.g. HIV infection and systemic lupus erythematosus, disorders  
XX of the endocrine system, e.g. diabetes mellitus, central nervous system  
XX disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral  
XX sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,  
XX parasitic and viral diseases. Accordingly, they exhibit many various  
XX activities including cytostatic, anticoagulant, thrombolytic,  
XX antianaemic, cardiant, vasotropic, antiasthmatic, antiinflammatory,  
XX nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,  
XX dermatological, antirheumatic, antiarthritic, antidiabetic,  
XX anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,  
XX fungicide, antiparasitic and virucidal, such that these polynucleotides  
XX can be used for gene therapy purposes and the development of appropriate  
XX vaccines. This polypeptide is a human protein of the invention.  
XX  
SQ Sequence 219 AA;

Query Match 100.0%; Score 861; DB 8; Length 219;  
Best Local Similarity 100.0%; Pred. No. 1.8e-82;  
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGLEGSARHVVFGS 60  
DB 57 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGLEGSARHVVFGS 116  
QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOQTEDSWVPASPDGLDPLTVRRHVPAWVL 120  
DB 117 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOQTEDSWVPASPDGLDPLTVRRHVPAWVL 176  
QY 121 LSRDPLDPNECGYQPPGAPPGGLGMPSSCGPRRRRRDTRTS 163

DB 177 LSRDPLDPNECGYQPPGAPPGGLGMPSSCGPRRRRRDTRTS 219  
RESULT 4  
AA36104  
ID AAY36104 standard; protein; 163 AA.  
XX  
XX AAY36104;  
XX  
DT 13-SEP-1999 (first entry)  
XX  
XX Extended human secreted protein sequence, SEQ ID NO. 489.  
XX  
XX Secreted protein; human; cytokine; cellular proliferation; cell movement;  
XX cellular differentiation; immune system regulator; anti-inflammatory;  
XX haematopoiesis regulator; tissue growth regulator; tumour inhibitor;  
XX reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;  
XX genetic disease.  
XX  
XX Homo sapiens.  
XX  
XX WO9931236-A2.  
XX  
XX 24-JUN-1999.  
XX  
XX 17-DEC-1998; 98WO-IB002122.  
XX  
XX 17-DEC-1997; 97US-0069957P.  
XX 09-FEB-1998; 98US-0074121P.  
XX 13-APR-1998; 98US-0081563P.  
XX 10-AUG-1998; 98US-0096116P.  
XX  
XX (GEST ) GENSET.  
XX  
XX Bougueleret L, Duclert A, Dumas Milne Edwards J;  
XX  
XX WPI; 1999-385906/32.  
XX N-PSDB; AAX97788.  
XX  
XX New isolated human secreted proteins.  
XX  
XX Claim 9; Page 423-424; 516pp; English.  
XX  
XX This sequence is encoded by an extended human secreted protein coding  
XX sequence of the invention. The secreted proteins can be used in treating  
XX or controlling a variety of human conditions. The secreted proteins may  
XX act as cytokines or may affect cellular proliferation or differentiation  
XX or may act as immune system regulators, haematopoiesis regulators, tissue  
XX growth regulators, regulators of reproductive hormones or cell movement  
XX or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or  
XX tumour inhibition activity. The DNAs can be used in forensic procedures  
XX to identify individuals or in diagnostic procedures to identify  
XX individuals having genetic diseases resulting from abnormal expression of  
XX the genes corresponding to the extended cDNAs. They are also useful for  
XX constructing a high resolution map of the human chromosomes. They can  
XX also be used for gene therapy to control or treat genetic diseases  
XX  
XX  
SQ Sequence 163 AA;

Query Match 94.4%; Score 813; DB 2; Length 163;  
Best Local Similarity 94.5%; Pred. No. 1.4e-77;  
Matches 154; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGLEGSARHVVFGS 60  
DB 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGLEGSARHVVFGS 60  
QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOQTEDSWVPASPDGLDPLTVRRHVPAWVL 120  
DB 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOQTEDSWVPASPDGLDPLTVRRHVPAWVL 120  
QY 121 LSRDPLDPNECGYQPPGAPPGGLGMPSSCGPRRRRRDTRTS 163

db 61 GRAAGKAVSCAETVKRRVPGI.HQI.TKI.XFI.OTENSWWVPXSPDTGLXPI.TVBRHYRAYWVI.120



Qy 61 GRAAGKAVSCAEIVKRRVP 79  
 |||||  
 Db 61 GRAAGKAVSCAEIVKRRVP 79

RESULT 7  
 AAG03599  
 ID AAG03599 standard; protein; 78 AA.  
 XX  
 AC AAG03599;  
 XX  
 DT 06-OCT-2000 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID NO: 7680.  
 XX  
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EPI033401-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 21-FEB-2000; 2000EP-00200610.  
 XX  
 PR 26-FEB-1999; 99US-0122487P.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Dumas Milne Edwards J, Duclert A, Giordano J;  
 DR WPI; 2000-500381/45.  
 DR N-PSDB; AAC03605.  
 XX  
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 XX  
 PS Claim 13; SEQ ID NO 7680; 71pp + Sequence Listing; English.  
 XX

CC The present sequence is a polypeptide encoded by one of a large number of  
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different  
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT  
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
 CC sequences derived from the 5' ends of mRNAs and even in those cases where  
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely  
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can  
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
 CC are also used in diagnostic, forensic, gene therapy and chromosome  
 CC mapping procedures. They are used to obtain upstream regulatory sequences  
 CC and to design expression and secretion vectors  
 XX  
 SQ Sequence 78 AA;  
 Query Match 45.9%; Score 395; DB 3; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-34;  
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEHYRKAGSVLPAPSPMPQLPPTLEMRVDRGSKIRNLGLALGRLEGGSARHVVFGS 60  
 |||||  
 Db 1 MEHYRKAGSVLPAPSPMPQLPPTLEMRVDRGSKIRNLGLALGRLEGGSARHVVFGS 60

Qy 61 GRAAGKAVSCAEIVKRRVP 78  
 |||||  
 Db 61 GRAAGKAVSCAEIVKRRVP 78

RESULT 8  
 AAM38973

ID AAM38973 standard; protein; 211 AA.  
 XX  
 AC AAM38973;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 2118.  
 XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI58129.  
 XX

PT Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX  
 PS Example 4; SEQ ID NO 2118; 10078pp; English.  
 XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX  
 SQ Sequence 211 AA;  
 Query Match 39.4%; Score 339; DB 4; Length 211;  
 Best Local Similarity 41.9%; Pred. No. 2.5e-27;  
 Matches 72; Conservative 26; Mismatches 56; Indels 18; Gaps 3;

Qy 1 MEHYRKAGSVLPAPSPMPQLPPTLEMRVDRGSKIRNLGLALGRLEG 50  
 |||||  
 Db 13 MENFRKVRSEAPRGCGAEGGPGFPADLAPGAVHMRVKEGSKIRNLMAFATASMAQP 72



PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 08-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147322P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 23.9%; Score 206; DB 3; Length 350;  
Best Local Similarity 29.9%; Pred. No. 5.4e-13;  
Matches 53; Conservative 23; Mismatches 65; Indels 36; Gaps 4;  
Qy- 1 MEHYRKAGSVLPAPSPMPQLPBDTLEMEVRDGSKTRNLLGLALGLEGSARHVFSGS 60  
Db 1 MDKIYQR-----VKPKADTPIDANEIRITSGGRARNYITYAMTLLQDKGSTEIVFKAM 53  
Qy 61 GRAAGKAVSCAETVKRRVPGLHQLTKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAWVL 120



PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151085P.  
PR 27-AUG-1999; 99US-0151086P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 01-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161931P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 23.9%; Score 206; DB 3; Length 350;

Best Local Similarity 29.9%; Pred. No. 5.4e-13;

Matches 53; Conservative 23; Mismatches 65; Indels 36; Gaps 4;

Qy 1 MEHYRKAGVELPAPSPMPQLPPTLEMYRVDGSKIRNLLGLALGRLEGGSARHVPFSGS 60

Db 1 MDKYQR-----VVKPKADTPIDANEIRITSQGRAENYITYAMTLIQDKGSTEVVFKAM 53

Qy 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTBDSWVPASPDGLDPLTVRRHVPAAWVL 120  
Db 54 GRAINKTVITVELIKRIPDLHQWTSIGSTDITDTWEPT--EEGLLPLETRHVSMTIT 111  
Qy 121 LSRDPLDPNECGYQPP-----GAPPGIGSMFSSSCGPRSRRAR 159  
Db 112 LSKIELNTSSVGQCPIPIELVKMGDIDYEGREGSPGGRG-----RGRGR 159

RESULT 12  
AAG51847  
ID AAG51847 standard; protein; 350 AA.  
XX AAG51847;  
XX AC  
XX 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 65844.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
XX  
PN EF1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 05-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.





[illegible]





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 10:10:14 ; Search time 390 Seconds  
(without alignments)

6622.347 Million cell updates/sec

Title: US-10-057-813-13

Perfect score: 492

Sequence: 1 atggagcactaccggaagc.....ctcgagacaccgcatcgtga 492

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_23Sep04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	492	100.0	660	12	ADM91197 Human DNA
2	492	100.0	660	12	ADM90779 Human DNA
3	492	100.0	894	5	Aah64942 Human sec
4	492	100.0	1089	8	Abx05092 Human nov
5	484	98.4	884	2	Aax97788 Extended
6	484	98.4	884	12	Adp19055 Human sec
7	385.6	78.4	1036	3	Aac95569 Human sec
8	385.6	78.4	1036	8	Ada98058 Human sec
9	385.6	78.4	1036	8	Ada43944 Human sec
10	385.6	78.4	1036	10	Adc20214 Human sec
11	385.6	78.4	1036	10	Adf10630 Human sec
12	373.8	76.0	469	9	Ach33887 Human end
13	368.4	74.9	1168	9	Adal1077 Human cDN
14	368.4	74.9	1168	10	Adj56254 Human cDN
15	236	48.0	303	3	Aac03605 Human sec
16	236	48.0	305	2	Aax40672 Human sec
17	223.8	45.5	560	10	Abt41218 Toxicity
18	78	15.9	1424	4	Aai59915 Human pol
19	74.6	15.2	1941	4	Aai58129 Human pol
20	74.6	15.2	1941	5	Adq98335 DNA encod
21	74.6	15.2	1941	9	Adb48095 Novel hum

22	74.6	15.2	2375	12	ADL12363	Adl12363 Human ste
23	68.2	13.9	412	8	ABX52110	Abx52110 Bovine ES
24	44.4	9.0	648	4	ABL15187	Abli15187 Drosophill
25	44.4	9.0	2000	8	ADA71938	Ada71938 Rice gene
26	41	8.3	2000	8	ADA71938	Ada71938 Rice gene
27	40.4	8.2	329	3	AAF15097	Aaf15097 Trichoder
28	39.6	8.0	2816	4	ABL15186	Abli15186 Drosophill
29	39.4	8.0	32189	4	AAL36926	Aal36926 Human mus
30	39.4	8.0	32189	5	ABA16341	Abal16341 Human ner
31	39.4	8.0	32189	5	ABA17492	Abal17492 Human ner
32	39.4	8.0	32189	5	ABA17790	Abal17790 Human ner
33	39.4	8.0	32189	8	ABX59914	Abx59914 cDNA enco
34	39.4	8.0	32189	12	ADJ30664	Adj30664 Human mus
35	37.2	7.6	1338	3	AAC43679	Aac43679 Zea maye
36	37.2	7.6	46819	9	ACD19059	Acd19059 E. coli 0
37	37.2	7.6	46897	10	ADC00585	Adc00585 Enterohae
38	36.6	7.4	3633	12	ADM66939	Adm66939 Murine ad
39	36.2	7.4	500	10	ADE81670	Adel81670 Arabidops
40	36.2	7.4	1053	12	ADN72182	Adn72182 Thale cre
41	36.2	7.4	1053	12	ADN72882	Adn72882 Thale cre
42	36.2	7.4	1071	3	AAC50707	Aac50707 Arabidops
43	36.2	7.4	1603	3	AAC50718	Aac50718 Arabidops
44	36.2	7.4	1605	3	AAC37507	Aac37507 Arabidops
45	36	7.3	11100	10	ADI23891	Adi23891 Streptomy

#### ALIGNMENTS

RESULT 1

ADM91197

ID ADM91197 standard; DNA; 660 BP.

AC ADM91197;

XX 03-JUN-2004 (first entry)

DE Human DNA encoding a pharmaceutically useful protein SeqID 590.

XX gene; ds; human; cancer; haematopoiesis; thrombosis; anaemia;

KW cardiovascular disorder; ischaemic heart disease;

KW acute myocardial infarction; respiratory disease; asthma; pneumonia;

KW gastric fibrosis; chronic renal failure; glomerulopathy;

KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;

KW HIV infection; sytemic lupus erythematosus; endocrine system;

KW diabetes mellitus; epilepsy; Alzheimer's disease;

KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;

KW fungal; parasitic; viral infection; cytostatic; anticoagulant;

KW thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;

KW antiinflammatory; nephrotropic; antiulcer; hepatotrophic;

KW immunosuppressive; antiallergic; dermatological; antirheumatic;

KW antithrptic; antidiabetic; anticonvulsant; neuroprotective; nootropic;

KW antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;

XX gene therapy; vaccine.

OS Homo sapiens.

PN WO2004020595-A2.

XX 11-MAR-2004.

PD 28-AUG-2003; 2003WO-US027107.

PF 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410951P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 XX  
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 PA (RIKEN-) RIKEN INST PHYSICAL & CHEM RES.  
 PA (DNAF-) DNAFORM KK.  
 XX  
 PI Williams LT, Chu K, Lee E, Hestir K;  
 XX  
 XX WPI; 2004-257410/24.  
 DR P-PSDB; ADM90988.  
 XX  
 PT New human polynucleotides and polypeptides, useful for diagnosing,  
 PT preventing and treating proliferative disorders, immune disorders,  
 PT cardiovascular disorders, or bacterial, fungal, parasitic and viral  
 PT diseases.  
 XX  
 XX Claim 1; SEQ ID NO 590; 254pp; English.  
 XX  
 XX This invention relates to novel isolated human polynucleotides and the  
 CC encoded proteins thereof. Specifically, it refers to proteases, kinases,  
 CC phosphatases, secreted and transmembrane proteins, as well as the derived  
 CC peptide fragments, which can be used to develop antibodies and screen for  
 CC small molecule agonists and antagonists that can modulate their  
 CC activities. The present invention describes polypeptides,  
 CC polynucleotides, vectors and host cells useful for diagnosing, preventing  
 CC and treating proliferative disorders, e.g. cancer, disorders of  
 CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,  
 CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory  
 CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the  
 CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,  
 CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune  
 CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders  
 CC of the endocrine system, e.g. diabetes mellitus, central nervous system  
 CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral  
 CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,  
 CC parasitic and viral diseases. Accordingly, they exhibit many various  
 CC activities including cytostatic, anticoagulant, thrombolytic,  
 CC antianaemic, cardiac, vasotropic, antiasthmatic, antiinflammatory,  
 CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,  
 CC dermatological, antirheumatic, antidiabetic, antidiarrhetic, antiparasitic,  
 CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,  
 CC fungicide, antiparasitic and virucidal, such that these polynucleotides  
 CC can be used for gene therapy purposes and the development of appropriate  
 CC vaccines. This polynucleotide is a human DNA sequence encoding a  
 CC polypeptide of the invention.  
 XX  
 XX Sequence 660 BP; 109 A; 229 C; 206 G; 116 T; 0 U; 0 Other;  
 SQ

Query Match  
 Best Local Similarity 100.0%; Score 492; DB 12; Length 660;  
 Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTGTAGAGCTCCAGCGCTTCCCAATGCCAG 60  
 DB 169 ATGGAGCACTACCGGAAAGCTGGCTGTAGAGCTCCAGCGCTTCCCAATGCCAG 228  
 QY 61 CTACCTCTGTATACCTCTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGTG 120  
 DB 229 CTACCTCTGTATACCTCTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGTG 288  
 QY 121 GGGTGGCTCTGGGTGGTGGAGGGCGGAGTGTGGAGTGTAGTGTCTCAGGTTCT 180  
 DB 289 GGGTGGCTCTGGGTGGTGGAGGGCGGAGTGTGGAGTGTAGTGTCTCAGGTTCT 348

QY 181 GCGAGGGCTGACGAGAAAGCTGTCAAGTGGCGTGTGAGATTGTCAAGCGGGGTCCAGGC 240  
 DB 349 GCGAGGGCTGACGAGAAAGCTGTCAAGTGGCGTGTGAGATTGTCAAGCGGGGTCCAGGC 408  
 QY 241 CTGCAACCAAGCTACCAAGCTAGTTTCTTTCAGATGAGGACAGCTGGGTCCAGCCTCA 300  
 DB 409 CTGCAACCAAGCTACCAAGCTAGTTTCTTTCAGATGAGGACAGCTGGGTCCAGCCTCA 468  
 QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGGCGCCCATGTGCTTGCAGTGTGGGTGCTG 360  
 DB 469 CCTGACACAGGGCTAGACCCCTCACAGTGGCGCCCATGTGCTTGCAGTGTGGGTGCTG 528  
 QY 361 CTCAGCGGGACCCCTCGACCCCAATGAGTGTGTTTACCAACCCCGAGGAGACCCCT 420  
 DB 529 CTCAGCGGGACCCCTCGACCCCAATGAGTGTGTTTACCAACCCCGAGGAGACCCCT 588  
 QY 421 GGCCTGGGTTCATGCCAGCTCCAGCTGGCGCTTCCCGAAGAGGCTCGAGAC 480  
 DB 589 GGCCTGGGTTCATGCCAGCTCCAGCTGGCGCTTCCCGAAGAGGCTCGAGAC 648  
 QY 481 ACCCGATCGTGA 492  
 DB 649 ACCCGATCGTGA 660

RESULT 2  
 ADM90779  
 ID ADM90779 standard; DNA; 660 BP.  
 AC ADM90779;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human DNA encoding a pharmaceutically useful protein SeqID 172.  
 XX  
 KW gene; ds; human; cancer; haematopoiesis; thrombosis; anaemia;  
 KW cardiovascular disorder; ischaemic heart disease;  
 KW acute myocardial infarction; respiratory disease; asthma; pneumonia;  
 KW cystic fibrosis; chronic renal failure; glomerulopathy;  
 KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;  
 KW HIV infection; systemic lupus erythematosus; endocrine system;  
 KW diabetes mellitus; epilepsy; Alzheimer's disease;  
 KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;  
 KW fungal; parasitic; viral infection; cytostatic; anticoagulant;  
 KW thrombolytic; antianaemic; cardiac; vasotropic; antiasthmatic;  
 KW antiinflammatory; nephrotropic; antidiabetic; antidiarrhetic;  
 KW immunosuppressive; antiallergic; dermatological; antirheumatic;  
 KW anticonvulsant; neuroprotective; nootropic; antipsoriatic; antibacterial;  
 KW fungicide; antiparasitic and virucidal; such that these polynucleotides  
 XX can be used for gene therapy purposes and the development of appropriate  
 XX vaccines. This polynucleotide is a human DNA sequence encoding a  
 OS polypeptide of the invention.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO2004020595-A2.  
 XX  
 XX 11-MAR-2004.  
 PD  
 XX  
 PF 28-AUG-2003; 2003WO-027107.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410951P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410962P.

PR	17-SEP-2002; 2002US-0411019P.	
PR	17-SEP-2002; 2002US-0411022P.	
PR	17-SEP-2002; 2002US-0411024P.	
PR	17-SEP-2002; 2002US-0411037P.	
PR	17-SEP-2002; 2002US-0411046P.	
PR	17-SEP-2002; 2002US-0411052P.	
PR	17-SEP-2002; 2002US-0411082P.	
PR	17-SEP-2002; 2002US-0411111P.	
XX		
(FIVE-) FIVE PRIME THERAPEUTICS INC.		
PA	(RIKE-) RIKEN INST PHYSICAL & CHEM RES.	
PA	(DNAF-) DNAFORM KK.	
XX		
PI	Williams LT, Chu K, Lee E, Hestir K;	
XX		
DR	WPI; 2004-257410/24.	
DR	P-PSDS; ADM90988.	
XX		
PT	New human polynucleotides and polypeptides, useful for diagnosing,	
PT	preventing and treating proliferative disorders, immune disorders,	
PT	cardiovascular disorders, or bacterial, fungal, parasitic and viral	
PT	diseases.	
XX		
PS	Claim 1; SEQ ID NO 172; 254pp; English.	
XX		
CC	This invention relates to novel isolated human polynucleotides and the	
CC	encoded proteins thereof. Specifically, it refers to proteases, kinases,	
CC	phosphatases, secreted and transmembrane proteins, as well as the derived	
CC	peptide fragments, which can be used to develop antibodies and screen for	
CC	small molecule agonists and antagonists that can modulate their	
CC	activities. The present invention describes polypeptides,	
CC	polynucleotides, vectors and host cells useful for diagnosing, preventing	
CC	and treating proliferative disorders, e.g. cancer, disorders of	
CC	haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,	
CC	e.g. ischaemic heart disease and acute myocardial infarction, respiratory	
CC	diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the	
CC	kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,	
CC	gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune	
CC	disorders, e.g. HIV infection and systemic lupus erythematosus, disorders	
CC	of the endocrine system, e.g. diabetes mellitus, central nervous system	
CC	disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral	
CC	sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,	
CC	parasitic and viral diseases. Accordingly, they exhibit many various	
CC	activities including cytostatic, anticoagulant, thrombolytic,	
CC	antianaemic, cardiact, vasotropic, antiasthmatic, antiinflammatory,	
CC	nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,	
CC	dermatological, antirheumatic, antiarthritic, antidiabetic,	
CC	anticonvulsant, neuroprotective, nootropic, antiparasitic, antibacterial,	
CC	fungicide, antiparasitic and virucidal, such that these polynucleotides	
CC	can be used for gene therapy purposes and the development of appropriate	
CC	vaccines. This polynucleotide is a human DNA sequence encoding a	
CC	polypeptide of the invention.	
XX		
SQ	Sequence 660 BP; 109 A; 229 C; 206 G; 116 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 492; DB 12; Length 660;	
	Best Local Similarity 100.0%; Pred. No. 5.9e-120;	
	Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 ATGAGGACACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60	
Db	169 ATGAGGACACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 228	
QY	61 CTACCTCCTGATACCTTGAGTCGGGTCCGAGATGCGACGACAAATTCGCAACTGTGCT 120	
Db	229 CTACCTCCTGATACCTTGAGATCGGGTCCGAGATGCGACGACAAATTCGCAACTGTGCT 288	
QY	121 GGGTTGGCTCTGGGTTCGGTTCGAGCGGCGGACGTGCTCGGCATGTAGTCTTCTCAGGTTCT 180	
Db	289 GGGTTGGCTCTGGGTTCGGTTCGAGCGGCGGACGTGCTCGGCATGTAGTCTTCTCAGGTTCT 348	
QY	181 GGCAAGGCTCGCAGGAAAGGCTGTTCAGTTCGCTGAGATTTGTCAGCGCGGGTCCCAAGC 240	

CC patients may be in need of restorative therapy. The GENSET polypeptides  
CC may be used as antigens in the production of antibodies and in assays to  
CC identify modulators (agonists and antagonists) of GENSET polypeptide  
CC expression and activity. The present sequence is a GENSET nucleic acid of  
CC the invention

XX SQ Sequence 894 BP; 189 A; 258 C; 255 G; 192 T; 0 U; 0 Other;  
Query Match 100.0%; Score 492; DB 5; Length 894;  
Best Local Similarity 100.0%; Pred. No. 6.3e-120;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGCACTACCGAAAGCTGCTCTGTAGAGTCCCGAGGCTTCCCAATGCCCCAG 60  
Db 78 ATGAGCACTACCGAAAGCTGCTCTGTAGAGTCCCGAGGCTTCCCAATGCCCCAG 137  
QY 61 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGCAGCAAAATTCGAACTGCTG 120  
Db 138 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGCAGCAAAATTCGAACTGCTG 197  
QY 121 GGGTTGGCTCTGGTCTGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
Db 198 GGGTTGGCTCTGGTCTGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 257  
QY 181 GGCAGGCTCAGGAAAGCTGTGAGTCTGGCTGAGATTGTCAAGGGCGGGTCCAGGC 240  
Db 258 GGCAGGCTCAGGAAAGCTGTGAGTCTGGCTGAGATTGTCAAGGGCGGGTCCAGGC 317  
QY 241 CTGACAGCTACCAAGCTAGTTTCTTTCAGACTGAGGAGAGTGGTCCAGGCTCA 300  
Db 318 CTGACAGCTACCAAGCTAGTTTCTTTCAGACTGAGGAGAGTGGTCCAGGCTCA 377  
QY 301 CCTGACACAGGCTAGACCCCTCAGAGTGGCGGCGGATGTGCTGAGTGTGGTCTG 360  
Db 378 CCTGACACAGGCTAGACCCCTCAGAGTGGCGGCGGATGTGCTGAGTGTGGTCTG 437  
QY 361 CTCAGCGGAGCCCTTGGAGCCCAATGATGTGTTTACCAACCCCGAGGAGACCCCT 420  
Db 438 CTCAGCGGAGCCCTTGGAGCCCAATGATGTGTTTACCAACCCCGAGGAGACCCCT 497  
QY 421 GGCTGGTTCATGCCAGCTCAGTGTGGCCCTCGTTTCCGAGAGAGGGTCCGAGAC 480  
Db 498 GGCTGGTTCATGCCAGCTCAGTGTGGCCCTCGTTTCCGAGAGAGGGTCCGAGAC 557  
QY 481 ACCCGATCGTA 492  
Db 558 ACCCGATCGTA 569

RESULT 4  
ABX05092  
ID ABX05092 standard; cDNA; 1089 BP.

XX AC ABX05092;

XX DT 17-JAN-2003 (first entry)

XX DE Human novel polynucleotide #107.

XX KW Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;  
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;  
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;  
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;  
KW fungal infection; bacterial infection; autoimmune disease; diabetes;  
KW atopic dermatitis.

XX OS Homo sapiens.

XX PN WO200274961-A1.

XX PD 26-SEP-2002.

XX PF 14-MAR-2002; 2002WO-US005109.

XX 15-MAR-2001; 2001US-00810173.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX WPI; 2003-040556/03.  
DR P-PSDB; ABU00014.  
XX New isolated polypeptides and polynucleotides, useful for preventing,  
PT treating or ameliorating medical conditions, such as cancer,  
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative  
PT disorders, and infections.  
XX Claim 1; SEQ ID NO 107; 235pp; English.  
XX The invention relates to human polynucleotides and the polypeptides they  
CC encode. The polynucleotides and polypeptides are useful in diagnostics,  
CC forensics, gene mapping, medical imaging, identification of mutations,  
CC responsible for genetic disorders or other traits, assessing biodiversity  
CC and producing many other types of data and products dependent on DNA and  
CC amino acid sequences. They are also useful for preventing, treating or  
CC ameliorating medical conditions, such as cancer, neurodegenerative  
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell  
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,  
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or  
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).  
CC Sequences ABX04986-ABX05511 represent human polynucleotides of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification but is based on sequence information supplied  
CC by the European Patent Office  
XX SQ Sequence 1089 BP; 221 A; 306 C; 331 G; 231 T; 0 U; 0 Other;

Query Match 100.0%; Score 492; DB 8; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 6.6e-120;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGTCCCGAGGCTTCCCAATGCCCCAG 60  
Db 275 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGTCCCGAGGCTTCCCAATGCCCCAG 334  
QY 61 CTACCTCTGATACCTTGTAGATGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 335 CTACCTCTGATACCTTGTAGATGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 394  
QY 121 GGGTTGGCTCTGGTCTGGTTGGAGGGCGGAGTGTGCTGGCATGTAGTGTCTCAGTTCT 180  
Db 395 GGGTTGGCTCTGGTCTGGTTGGAGGGCGGAGTGTGCTGGCATGTAGTGTCTCAGTTCT 454  
QY 181 GGCAGGCTGCAGGAAAGCTGTGAGTGTGCTGAGATTGTCAAGCGGGTCCAGGC 240  
Db 455 GGCAGGCTGCAGGAAAGCTGTGAGTGTGCTGAGATTGTCAAGCGGGTCCAGGC 514  
QY 241 CTGACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGTGGTCCAGCCTCA 300  
Db 515 CTGACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGTGGTCCAGCCTCA 574  
QY 301 CTGACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGTGGTCCAGCCTCA 360  
Db 575 CTGACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGTGGTCCAGCCTCA 634  
QY 361 CTGACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGTGGTCCAGCCTCA 420  
Db 635 CTGACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGTGGTCCAGCCTCA 694  
QY 421 GGCTGGTTCATGCCAGCTCAGTGTGGCCCTCGTTTCCGAGAGAGGGTCCGAGAC 480  
Db 695 GGCTGGTTCATGCCAGCTCAGTGTGGCCCTCGTTTCCGAGAGAGGGTCCGAGAC 754

```

QY 481 ACCCGATCGTGA 492
Db 755 ACCCGATCGTGA 766

RESULT 5
AA97788
ID AAX97788 standard; DNA; 884 BP.
AC AAX97788;
XX
XX
DT 13-SEP-1999 (first entry)
XX
XX Extended human secreted protein coding sequence, SEQ ID NO. 353.
XX
XX Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease; 88.
XX
XX Homo sapiens.
OS
XX
XX WO9931236-A2.
PN
XX
XX 24-JUN-1999.
XX
XX 17-DEC-1998; 98WO-IB002122.
XX
XX 17-DEC-1997; 97US-0069957P.
PR 09-FEB-1998; 98US-0074121P.
PR 13-APR-1998; 98US-0081563P.
PR 10-AUG-1998; 98US-0096116P.
XX
XX (GEST ) GENSET.
PA
XX
XX Bougueleret L, Duclert A, Dumas Milne Edwards J;
PI
XX
XX WPI; 1999-385906/32.
DR
XX P-PSDB; AAY36104.
DR
XX
XX New isolated human secreted proteins.
PT
XX
XX Claim 1; Page 423-424; 516pp; English.
PS
XX
XX This sequence represents an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases
XX
XX Sequence 884 BP; 187 A; 254 C; 238 G; 190 T; 0 U; 15 Other;
SQ

Query Match 98.4%; Score 484; DB 2; Length 884;
Best Local Similarity 97.8%; Pred. No. 8.1e-118;
Matches 481; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
Db 69 ATGGAGCACTACCGGAAAGTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 128
QY 61 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGGAGGAGGAAATTCGCACTGCTG 120
Db 129 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGGAGGAGGAAATTCGCACTGCTG 188

```

CC human tissues and cells from non-human tissues and cells, and for  
 CC distinguishing between human tissues and cells that do or do not express  
 CC the polynucleotides comprising the cDNAs. The polynucleotides and  
 CC polypeptides are useful in forensic procedures or diagnostic procedures  
 CC to identify individuals with genetic diseases resulting from abnormal  
 CC expression of the genes corresponding to the cDNAs. The sequences are  
 CC also useful in gene therapy to control or treat genetic diseases. This  
 CC sequence represents a human secreted polynucleotide of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html.

XX Sequence 884 BP; 187 A; 254 C; 238 G; 190 T; 0 U; 15 Other;

Query Match 98.4%; Score 484; DB 12; Length 884;

Best Local Similarity 97.8%; Pred. No. 8.1e-118;

Matches 481; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCCTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
 DB 69 ATGGAGCCTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 128  
 QY 61 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGCGAGCAAAATTCGCAACCTGCTG 120  
 DB 129 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGCGAGCAAAATTCGCAACCTGCTG 188  
 QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCCGATCTGCGATCTAGTGTCTCAGGTTCT 180  
 DB 189 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCCGATCTAGTGTCTCAGGTTCT 248  
 QY 181 GGCAGGGCTGCGGAAAGGCTGTGAGTGGCTGTGAGATTGTCAAGCGGGGGTCCAGGC 240  
 DB 249 GGCAGGGCTGCGGAAAGGCTGTGAGTGGCTGTGAGATTGTCAAGCGGGGGTCCAGGC 308  
 QY 241 CTCACAGCTACCAAGCTACGTTCTTCAGACTGAGACAGCTGGGTCCAGCTCA 300  
 DB 309 CTCACAGCTACCAAGCTACGTTCTTCAGACTGAGACAGCTGGGTCCAGCTCA 368  
 QY 301 CTGACACAGGGCTAGACCCCTTCAGTGCAGTCCGCGCCATGTGCTGCTGCTGGTGGTGG 360  
 DB 369 CTGACACAGGGCTARACCCCTTCAGTGCAGTCCGCGCCATGTGCTGCTGCTGGTGGTGG 428  
 QY 361 CTCAGCGGGAGCCCTTGGACCCCAATGATGTGTTTACCAACCCCGAGGACACCCCT 420  
 DB 429 CTCAGCGGGAGCCCTTGGACCCCAATGATGTGTTTACCAACCCCGAGGACACCCCT 488  
 QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCGAGAAAGGGCTGAGAC 480  
 DB 489 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCGAGAAAGGGCTGAGAC 548  
 QY 481 ACCGATCGTGA 492  
 DB 549 ACCGATCGTGA 560

RESULT 7

AAC95569

ID AAC95569 standard; cDNA; 1036 BP.

XX AAC95569;

AC AAC95569;

XX 21-FEB-2001 (first entry)

DE Human secreted protein gene 49 SEQ ID NO:59.

XX Human; secreted protein; cytostatic; immunosuppressive; neurotropic;

XX neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;

XX antinflammatory; antitumor; antiparasitic; cardiac; cancer; immune disease; allergy;

XX cardiovascular disorder; wound healing; infection; neurological disease;

XX ss.

XX Homo sapiens.

OS

XX WO200061596-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US008983.

XX 09-APR-1999; 99US-0128703P.

XX 14-JAN-2000; 2000US-0176068P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM, Komatsoulis G;

XX WPI; 2000-611865/58.

XX P-PSDB; AAB52060.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful in

XX the prevention, treatment and diagnosis of cancer, immune disorders,

XX cardiovascular disorders and neurological diseases.

XX Claim 1; Page 446; 505pp; English.

XX Polynucleotide sequences AAC95521 - AAC95570 represent cDNA encoding 50

XX human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -

XX AAB52103 represent alternative polypeptides encoded by the genes and

XX amino acid sequences with which they share homology. The genes and

XX proteins have activities dependent on the tissues and cells in which they

XX are expressed. Examples of their activities include cytostatic;

XX immunosuppressive; neurotropic; neuroprotective; antiviral; antiallergic;

XX hepatotropic; antidiabetic; antinflammatory; antitumor; antiparasitic;

XX anticonvulsant; antifungal; antiparasitic; antitumor; antidiabetic;

XX The secreted proteins, polypeptides, antagonists and agonists may be

XX useful in treating, preventing and/or diagnosing diseases and disorders

XX such as cancer, particularly breast and ovarian cancer, and other cancers

XX of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,

XX liver, lung, or urogenital. Immune disorders such as Addison's disease,

XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,

XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid

XX arthritis and ulcerative colitis; cardiovascular disorders such as

XX myocardial ischaemia; wound healing; neurological diseases such as

XX cerebral anoxia and epilepsy; and infectious diseases such as viral,

XX bacterial, fungal and parasitic infections may also be treated using the

XX proteins and polynucleotides of the invention. Sequences AAC95512 -

XX AAC95520 and AAB52011 are used in the isolation and characterisation of

XX the proteins and polynucleotides of the invention

XX SQ Sequence 1036 BP; 368 A; 238 C; 235 G; 192 T; 0 U; 3 Other;

Query Match 78.4%; Score 385.6; DB 3; Length 1036;

Best Local Similarity 87.8%; Pred. No. 7.9e-92;

Matches 432; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

QY 1 ATGAGCCTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
 DB 44 ATGAGCCTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 103  
 QY 61 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGCGAGCAAAATTCGCAACCTGCTG 120  
 DB 104 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGCGAGCAAAATTCGCAACCTGCTG 163  
 QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCCGATCTGCGATGTAGTGTCTCAGGTTCT 180  
 DB 164 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCCGATCTGCGATGTAGTGTCTCAGGTTCT 222  
 QY 181 GGCAGGGCTGCGGAAAGGCTGTGAGTGGCTGTGAGATTGTCAAGCGGGGGTCCAGGC 240  
 DB 223 GGTGGGCTGCGGAAAGGCTGTGAGTGGCTGTGAGATTGTCAAGCGGGGGTGTACCGGC 282  
 QY 241 CTGACAGCTACCAAGCTACGTTTCTTCAAGCTGAGGACAGCTGGTCCAGCTCA 300  
 DB 283 CTGACAGCTTACCAAGCTTCTTCTCTGAGACCGAGGACAGCTGGTCCAGCTCA 342



QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGGCGCGCCCATGTGCTGCAGTGTGGTCTG 360  
DB 343 CCTGACACAGCGCTGGATCCCTCACGGTACGCGCCCATGTACCTGCAGTGTGGTACTG 402  
QY 361 CTACGCGGACCCCTGGACCCCAATGAGTGTGTTACCAACCCCGAGGACCCCT 420  
DB 403 CTACGCGGACCCCTGGACCCCAATGAGTGTGCTTACCAAGCTCCAGAGACCCCT 462  
QY 421 GGCTGGTTCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480  
DB 463 GGCTGGGCCCCACATCAAGCTCCAGCTGTGTTCCAGACCCCGAAGAGGGTTCGAGAC 522  
QY 481 ACCGATCGTGA 492  
DB 523 ACCTGGTCTTAA 534

## RESULT 8

ADA98058  
ID ADA98058 standard; cDNA; 1036 BP.

XX AC  
ADA98058;

DT 20-NOV-2003 (first entry)

XX DE  
Human secreted protein cDNA sequence #152.

XX KW human; secreted protein; cardiovascular disorder; arrhythmia;  
KW atherosclerosis; stroke; endocarditis; congestive heart failure;  
KW rheumatic heart disease; cardiomyopathy; haemorrhoids; varicose veins;  
KW migraine; thrombosis; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; renal disorder; proliferative disorder; cancer; gene;  
KW ss.

XX OS  
Homo sapiens.

XX PN  
WO2003004623-A2.

XX PD  
16-JAN-2003.

XX PF  
26-MAR-2002; 2002WO-US0009922.

XX PR  
27-MAR-2001; 2001US-0278650P.

XX PR  
12-SEP-2001; 2001US-00950082.

XX PR  
12-SEP-2001; 2001US-00950083.

XX PA  
(HUMA-) HUMAN GENOME SCI INC.

XX PI  
Rosen CA, Ruben SM;

XX PI  
WPI; 2003-247946/24.

XX PT New human secreted polypeptide and nucleic acid molecules, useful for  
PT diagnosing, preventing, prognosticating or treating cardiovascular  
PT disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or  
PT thrombosis).

XX PS  
Claim 1; SEQ ID NO 162; 1572pp; English.

XX CC The invention comprises the amino acid and coding sequence of human  
CC secreted proteins. The DNA and protein sequences of the invention are  
CC useful in the treatment of cardiovascular disorders, such as: arrhythmia,  
CC atherosclerosis, stroke, endocarditis, congestive heart failure,  
CC rheumatic heart disease, cardiomyopathy, haemorrhoids, varicose veins,  
CC migraine, or thrombosis. The DNA and protein sequences may also be used  
CC for treating or preventing: neural disorders, immune system disorders,  
CC muscular disorders, reproductive disorders, gastrointestinal disorders,  
CC pulmonary disorders, renal disorders, proliferative disorders and/or  
CC cancerous diseases. The present cDNA sequence encodes a human secreted  
CC protein of the invention. NOTE: The present sequence is shown on the WIPO  
CC website.

XX SQ Sequence 1036 BP; 368 A; 238 C; 235 G; 192 T; 0 U; 3 Other;  
Query Match 78.4%; Score 385.6; DB 8; Length 1036;  
Best Local Similarity 87.8%; Pred. No. 7.9e-92;  
Matches 432; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

QY 1 ATGAGACACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCGCAG 60  
DB 44 ATGAGACACTACCGGAAAGCTGGCTCTGTGGAACCTCCAGCACCTTCTCCGATGCCGCG 103  
QY 61 CTACCTCTCTGATACCTTTCAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
DB 104 CTACCTCTCTGATACCTTTCAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 163  
QY 121 GGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGTCTCGGATGTAGTGTCTCAGGTTCT 180  
DB 164 GGGCTGGCACTGGTTCGATTGGA-GGTGGTGGTGCACGGCATGTGGTGTCTCAGGTTCT 222  
QY 181 GGCAGGGCTGCAGGAAAGCTGTTCAGCTGCGCTGAGATTGTCAAGCGCGGGTCCCGAGGC 240  
DB 223 GGTGGGGCTGCAGGAAAGCTGTTCAGCTGCGCTGAGATTGTCAAGCGGGTGTACCGGGC 282  
QY 241 CTGCACACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCGAGCTCA 300  
DB 283 CTGCACACAGCTTACCAAGCTGCGCTTCTTCAGAGCCGAGCAGCTGGGTGCGAGTCTCA 342  
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGGCGCGCCCATGTGCTGCAGTGTGGTCTG 360  
DB 343 CCTGACACAGGGCTGGATCCCTCACGGTACGCGCCCATGTACTCTGAGTGTGGTACTG 402  
QY 361 CTACGCGGAGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACCCCT 420  
DB 403 CTACGCGGAGACCCCTGGACCCCAATGAGTGTGGTACCAGGCTCCAGGAGCACCCCT 462  
QY 421 GGCTGGTTCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGTTCGAGAC 480  
DB 463 GGCTGGGCCCCACATCAAGCTCCAGCTGTGGTCCAGACCCCGAAGAGGGTTCGAGAC 522  
QY 481 ACCGATCGTGA 492  
DB 523 ACCTGGTCTTAA 534

## RESULT 9

ADA43944  
ID ADA43944 standard; cDNA; 1036 BP.

XX AC  
ADA43944;

XX DT  
20-NOV-2003 (first entry)

XX DE  
Human secreted protein cDNA SEQ ID 132.

XX KW Gene therapy; human; Antidiabetic; Anorectic; Ophthalmological;  
KW Neuroprotective; Cerebroprotective; Antianemic; gene; ss.

XX OS  
Homo sapiens.

XX PN  
WO2003000865-A2.

XX PD  
03-JAN-2003.

XX PF  
26-MAR-2002; 2002WO-US0009105.

XX PR  
27-MAR-2001; 2001US-0278650P.

XX PR  
12-SEP-2001; 2001US-00950082.

XX PR  
12-SEP-2001; 2001US-00950083.

XX PA  
(HUMA-) HUMAN GENOME SCI INC.

XX PI  
Rosen CA, Ruben SM;

XX PI  
XX





```
XX PR 30-JUL-2001; 2001US-00918995.
XX PA (DRMA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX DR WPI; 2003-615964/58.
XX PT New polynucleotide sequences obtained from various cDNA libraries, useful
XX PT as hybridization probes, as oligomers for PCR, for chromosome and gene
XX PT mapping, in the recombinant production of protein, or in generating
XX PT antisense DNA or RNA.
XX PS Claim 1; SEQ ID NO 21099; 44pp; English.
XX CC The invention relates to an isolated polynucleotide comprising any one of
XX CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX CC determined by the technique of SBH (sequencing by hybridisation). Also
XX CC included is a purified polypeptide comprising a sequence corresponding to
XX CC a reading frame of the novel polynucleotide. The nucleic acid sequences
XX CC are useful in diagnostics as expressed sequence tags (EST) for
XX CC identifying expressed genes or for physical mapping of the human genome,
XX CC in forensics, in assessing biodiversity, or in identifying mutations
XX CC responsible for genetic disorders and other traits. The nucleotide
XX CC sequences are also useful as hybridisation probes, as oligomers for PCR,
XX CC for chromosome and gene mapping, in the recombinant production of
XX CC protein, or in generating antisense DNA or RNA. The purified polypeptide
XX CC is useful for generating antibodies specific for it. The present sequence
XX CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX CC for this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX SQ Sequence 469 BP; 93 A; 141 C; 149 G; 86 T; 0 U; 0 Other;
Query Match 76.0%; Score 373.8; DB 9; Length 469;
Best Local Similarity 99.2%; Pred. No. 8.5e-89;
Matches 386; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 ATGGAGCAGCTACCGGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 60
DB 81 ATGGAGCAGCTACCGGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 140
QY 61 CTACCTCCTGATACCTTCCAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
DB 141 CTACCTCCTGATACCTTCCAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 200
QY 121 GGGTTGGCTTGGGTCGGTTGGAGGGCGGAGTGCTCGGATGTAGTGTCTCAG-GTTC 179
DB 201 GGGTTGGCTTGGGTCGGTTGGAGGGCGGAGTGCTCGGATGTAGTGTCTCAGAGTTC 260
QY 180 TGCGAGGGCTGCAGGAAAGCTGTGAGCTCGGCTGAGATTGTCAGCGGGTCCAGG 239
DB 261 TGCGAGGGCTGCAGGAAAGCTGTGAGCTCGGCTGAGATTGTCAGCGGGTCCAGG 320
QY 240 CTGCAACCAAGCTACCAAGCTAGTTTCTTCCATGAGAGTGGAGCAGCTGGGTCCAGCCCTC 299
DB 321 CTGCAACCAAGCTACCAAGCTAGTTTCTTCCATGAGAGTGGAGCAGCTGGGTCCAGCCCTC 380
QY 300 ACTTGACACAGGGCTAGACCCCTCAGTGCGGCCGCCATGTGCTGCAAGTGTGGGTGCT 359
DB 381 ACTTGACACAGGGCTAGACCCCTCAGTGCGGCCGCCATGTGCTGCAAGTGTGGGTGCT 440
QY 360 GCTCAGCGGGACCCCTCGACCCCAATG 388
DB 441 GCTCATCCGGGACCCCTCGACCCCAATG 469
```

---

```
RESULT 13
ADAL1077/C
ID ADAL1077 standard; cDNA; 1168 BP.
XX AC ADAL1077;
XX DT 06-NOV-2003 (first entry)
XX DE Human cDNA differentially expressed in colon cancer #1.
XX KW ss; differential expression; colon cancer; cancer; human.
XX OS Homo sapiens.
XX PN US2002160382-A1.
XX PD 31-OCT-2002.
XX PF 11-OCT-2001; 2001US-00981353.
XX PR 11-OCT-2000; 2000US-0239841P.
XX PA (LASE/) LASEK A W.
XX PA (JONE/) JONES D A.
XX PI Lasek AW, Jones DA;
XX DR WPI; 2003-265756/26.
XX CC New combination comprising cDNAs that are differentially expressed in
XX CC colon disorder, useful for diagnosing, treating, staging or monitoring
XX CC treatment for colon cancers.
XX PS Claim 1; SEQ ID NO 1; 231pp; English.
XX CC The invention relates to a combination comprising cDNAs that are
XX CC differentially expressed in colon disorder. The methods and compositions
XX CC of the present invention are useful for diagnosing, treating, staging or
XX CC monitoring treatment for colon cancer. They are also useful in high
XX CC throughput methods for using cDNAs to detect differential expression of
XX CC nucleic acids in a sample, screening molecules or compounds to identify a
XX CC ligand which specifically binds a cDNA and using a protein to screen
XX CC molecules or compounds to identify at least one ligand which specifically
XX CC binds the protein. The present sequence represents a human cDNA
XX CC differentially expressed in colon cancer.
XX SQ Sequence 1168 BP; 249 A; 346 C; 339 G; 234 T; 0 U; 0 Other;
Query Match 74.9%; Score 368.4; DB 9; Length 1168;
Best Local Similarity 99.7%; Pred. No. 2.8e-87;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 123 GTTGCTCTGGGTCGGTTGGAGGGCGGAGTGCTCGGATGTAGTGTCTCAGGTTCTGG 182
DB 871 GCTGGCTCTGGGTCGGTTGGAGGGCGGAGTGCTCGGATGTAGTGTCTCAGGTTCTGG 812
QY 183 CAGGCTCGAGAAAGGCTGTGAGTGGCTGAGATTGCAAGCGCGGGTCCAGGCT 242
DB 811 CAGGCTCGAGAAAGGCTGTGAGTGGCTGAGATTGCAAGCGCGGGTCCAGGCT 752
QY 243 GCACAGCTCACCAGCTACGTTTCCCTTCAGATGAGGACAGCTGGGTCCAGGCTCACC 302
DB 751 GCACAGCTCACCAGCTACGTTTCCCTTCAGATGAGGACAGCTGGGTCCAGGCTCACC 692
QY 303 TGACACAGGCTAGACCCCTCCAGATGCGCGCCGATGTGCTGAGTGTGGGTGCTGCT 362
DB 691 TGACACAGGCTAGACCCCTCCAGATGCGCGCCGATGTGCTGAGTGTGGGTGCTGCT 632
QY 363 CAGCGGGACCCCTCGACCCCAATGATGTGGTTACCAACCCCGAGAGCACCCCTGG 422
DB 631 CAGCGGGACCCCTCGACCCCAATGATGTGGTTACCAACCCCGAGAGCACCCCTGG 572
QY 423 CCTGGGTTCCATGCCAGCTCCAGCTGCGGCCCTCGTTCCCGAAGAGGGGCTCGAGACAC 482
```

Db 571 CCTGGGTTCATGCCAGCTCCAGCTGTGCGCCCTGTTCCGAGAGAGGCTCGAGACAC 512  
QY 483 CCGATCGTGA 492  
Db 511 CCGATCGTGA 502

RESULT 14  
ADJ56254/c  
ID ADJ56254 standard; cDNA; 1168 BP.  
XX  
AC ADJ56254;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human cDNA differentially expressed in MYCN activated cells SeqID 60.  
XX  
KW human; differential expression; transactivator; proto-oncogene;  
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;  
KW MYCN activated cell.  
XX  
OS Homo sapiens.  
XX  
PN US2003119009-A1.  
XX  
PD 26-JUN-2003.  
XX  
PF 25-FEB-2002; 2002US-00084817.  
XX  
PR 23-FEB-2001; 2001US-0270784P.  
XX  
PA (STUA//) STUART S G.  
PA (NUCH//) NUCHTERN J G.  
PA (PLOW//) PLOW S E.  
PA (SHOH//) SHOHET J M.  
XX  
PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;  
XX  
WPI; 2003-635698/60.  
XX  
PT New genes regulated by MYCN activation, useful in gene therapy,  
PT particularly for treating a subject with e.g. neuroblastoma or other  
PT cancers, or for diagnosing, staging or monitoring the treatment of the  
PT cancer.  
XX  
PS Claim 1; SEQ ID NO 60; 27pp; English.  
XX  
CC This invention relates to novel isolated cDNAs that are differentially  
CC expressed in MYCN activated cells. Specifically, it refers to  
CC polynucleotide sequences that exhibit differential expression patterns in  
CC cells activated by the transactivator MYCN, where MYCN is a proto-  
CC oncogene that is amplified in neuroblastoma cells and is common in small  
CC cell lung cancers. The present invention describes these cDNA molecules  
CC as useful for in hybridisation assays to detect expression of nucleic  
CC acids (or complementary nucleic acids) in a present in a given sample, as  
CC well as for screening assays by identifying molecules or compounds that  
CC specifically bind the cDNA as a ligand and modulate function or activity.  
CC Accordingly, these compositions exhibit cytostatic activity and can also  
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA  
CC that is differentially expressed in MYCN activated cells, given in an  
CC exemplification of the invention. NOTE: This sequence does not appear in  
CC the printed specification but has been obtained in electronic format from  
CC the US Patent Office at  
CC ftp.segdata.uspco.gov/sequence.html?DocID=20030119009.  
XX  
SQ Sequence 1168 BP; 249 A; 346 C; 339 G; 234 T; 0 U; 0 Other;  
Query Match 74.9%; Score 368.4; DB 10; Length 1168;  
Best Local Similarity 99.7%; Pred. No. 2.8e-87;  
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 123 GTTGGCTCTGGGTTCGTTGGAGGCGCGCAGTGTGCGCATGTAGTTCTCAGGTTCTGG 182

Db 871 GCTGGCTCTGGGTTCGTTGGAGGCGCGCAGTGTGCGCATGTAGTTCTCAGGTTCTGG 812  
QY 183 CAGGGCTGCAGGAAGAGCTGTGTCAGTGCCTGAGATTGTCAAGCGCGGTCCAGGCTT 242  
Db 811 CAGGGCTGCAGGAAGAGCTGTGTCAGTGCCTGAGATTGTCAAGCGCGGTCCAGGCTT 752  
QY 243 GCACACAGCTCACCAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCACC 302  
Db 751 GCACACAGCTCACCAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCACC 692  
QY 303 TGACACAGGCTAGACCCCTCAGTGCCTGAGTGCCTGAGTGCCTGAGTGCCTGAGTGCCT 362  
Db 691 TGACACAGGCTAGACCCCTCAGTGCCTGAGTGCCTGAGTGCCTGAGTGCCTGAGTGCCT 632  
QY 363 CAGCGGGAGCCCTCGGACCCCAATGAGTGTGTTACCAACCCCGAGGAGCAGCCCTTGG 422  
Db 631 CAGCGGGAGCCCTCGGACCCCAATGAGTGTGTTACCAACCCCGAGGAGCAGCCCTTGG 572  
QY 423 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCGCGAAGAGGCTCGAGACAC 482  
Db 571 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCGCGAAGAGGCTCGAGACAC 512  
QY 483 CCGATCGTGA 492  
Db 511 CCGATCGTGA 502

RESULT 15  
AAC03605  
ID AAC03605 standard; cDNA; 303 BP.  
XX  
AC AAC03605;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 3603.  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-00200610.  
XX  
PR 26-FEB-1999; 99US-0122487P.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duciert A, Giordano J;  
XX  
WPI; 2000-500381/45.  
XX  
P-PSDB; AAG03599.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
PS Claim 1; SEQ ID NO 3603; 71pp + Sequence Listing; English.  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. An ORF has been identified within the  
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
CC derived from 30 different tissues. EST sequences usually correspond  
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
CC often obtained from oligo-dr primed cDNA libraries. Such ESTs are not  
CC well suited for isolating cDNA sequences derived from the 5' ends of  
CC mRNAs and even in those cases where longer cDNA sequences have been  
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
CC gene therapy and chromosome mapping procedures. They are used to obtain  
CC upstream regulatory sequences and to design expression and secretion  
CC vectors

XX

SQ Sequence 303 BP; 57 A; 81 C; 105 G; 60 T; 0 U; 0 Other;

Query Match 48.0%; Score 236; DB 3; Length 303;  
Best Local Similarity 100.0%; Pred. NO. 1.8e-52;  
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
|||  
Db 68 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 127  
|||  
Qy 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
|||  
Db 128 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 187  
|||  
Qy 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
|||  
Db 188 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 247  
|||  
Qy 181 GGCAGGGCTGCAGGAAAGGCTGTTCAGCTGGCTGAGATTGTCAAGCGCGGGTCCC 236  
|||  
Db 248 GGCAGGGCTGCAGGAAAGGCTGTTCAGCTGGCTGAGATTGTCAAGCGCGGGTCCC 303  
|||

Search completed: January 31, 2005, 16:45:02  
Job time : 395 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:06:38 ; Search time 135.925 Seconds  
(without alignments)  
433.254 Million cell updates/sec

Title: US-10-057-813-14

Perfect score: 861

Sequence: 1 MEHYRKAGSVLPAPMPQ.....SMPSSCGPRRRRDTRTS 163

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	100.0	163	9	US-09-731-872-459
2	861	100.0	163	10	US-09-876-997-459
3	813	94.4	163	11	US-09-978-360A-668
4	305	35.4	72	17	US-10-425-115-237341
5	227	26.4	281	15	US-10-424-599-237821
6	224	26.0	242	17	US-10-425-115-254668
7	204.5	23.8	245	15	US-10-424-599-229595
8	203.5	23.6	276	17	US-10-425-115-308895
9	203.5	23.6	276	17	US-10-425-115-308929
10	203.5	23.6	278	15	US-10-425-114-60870
11	203.5	23.6	301	15	US-10-425-114-61107
12	203.5	23.6	315	15	US-10-425-114-54515
13	203.5	23.6	315	15	US-10-425-114-58679
					Sequence 459, App
					Sequence 668, App
					Sequence 237341,
					Sequence 237821,
					Sequence 254668,
					Sequence 229595,
					Sequence 308895,
					Sequence 308929,
					Sequence 60670, A
					Sequence 61107, A
					Sequence 54515, A
					Sequence 58679, A

14	203.5	23.6	325	15	US-10-425-114-60913	Sequence 60913, A
15	202	23.5	178	17	US-10-425-115-232929	Sequence 232929,
16	202	23.5	671	16	US-10-437-963-109378	Sequence 109378,
17	201.5	23.4	269	16	US-10-437-963-156973	Sequence 156973,
18	200.5	23.3	263	16	US-10-437-963-152741	Sequence 152741,
19	200	23.2	262	16	US-10-437-963-131078	Sequence 131078,
20	199.5	23.2	275	16	US-10-767-701-45240	Sequence 45240, A
21	198	23.0	294	15	US-10-425-114-60641	Sequence 60641, A
22	193.5	22.5	265	17	US-10-425-115-308917	Sequence 308917,
23	191	22.2	265	17	US-10-425-115-201476	Sequence 201476,
24	190.5	22.1	261	16	US-10-767-701-42504	Sequence 42504, A
25	188	21.8	287	16	US-10-437-963-132743	Sequence 132743,
26	188	21.8	359	17	US-10-425-115-231292	Sequence 231292,
27	187	21.7	261	16	US-10-437-963-184232	Sequence 184232,
28	185	21.5	320	17	US-10-739-930-10929	Sequence 10929, A
29	172	20.0	318	17	US-10-425-115-266917	Sequence 266917,
30	102.5	11.9	242	16	US-10-437-963-134549	Sequence 134549,
31	96	11.1	198	17	US-10-425-115-365780	Sequence 365780,
32	88.5	10.3	213	17	US-10-425-115-378508	Sequence 378508,
33	88	10.2	255	17	US-10-425-115-242952	Sequence 242952,
34	88	10.2	581	16	US-10-437-963-201766	Sequence 201766,
35	87.5	10.2	625	14	US-10-156-761-10693	Sequence 10693, A
36	86.5	10.0	174	14	US-10-094-749-2987	Sequence 2987, Ap
37	86.5	10.0	353	17	US-10-425-115-224520	Sequence 224520,
38	86.5	10.0	359	15	US-10-296-115-1274	Sequence 1274, Ap
39	86.5	10.0	999	16	US-10-437-963-181440	Sequence 181440,
40	86	10.0	186	17	US-10-425-115-220639	Sequence 220639,
41	85.5	9.9	639	16	US-10-408-765A-1054	Sequence 1054, Ap
42	85.5	9.9	774	17	US-10-723-860-1706	Sequence 1706, Ap
43	85	9.9	436	14	US-10-156-761-11365	Sequence 11365, A
44	84.5	9.8	234	15	US-10-108-260A-2464	Sequence 2464, Ap
45	84.5	9.8	283	16	US-10-437-963-137561	Sequence 137561,

ALIGNMENTS

RESULT 1

US-09-731-872-459  
; Sequence 459, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bouquelieret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 459  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-731-872-459

Query Match 100.0%; Score 861; DB 9; Length 163;  
Best Local Similarity 100.0%; Pred. No. 9.7e-77;  
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEHYRKAGSVLPAPMPQPPDTTLEMRVDRGSKIRNLLGLALGRLEGSGARHVVFGS	60
Db	1	MEHYRKAGSVLPAPMPQPPDTTLEMRVDRGSKIRNLLGLALGRLEGSGARHVVFGS	60
Qy	61	GRAAGKAVSCAEIVKRRVFGQLHQLTKLRFLOTEDSWVPASPDGLPLTVRRHVPAAVWL	120
Db	61	GRAAGKAVSCAEIVKRRVFGQLHQLTKLRFLOTEDSWVPASPDGLPLTVRRHVPAAVWL	120

Qy 121 LSRDPLDNECGYQPPGAPGLGMPSSSCGPRRRRRDTRS 163  
|  
Db 121 LSRDPLDNECGYQPPGAPGLGMPSSSCGPRRRRRDTRS 163

## RESULT 2

US-09-876-997-459  
; Sequence 459, Application US/09876997  
; Publication No. US2003015292A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78 USA CIP  
; CURRENT APPLICATION NUMBER: US/09/876,997  
; CURRENT FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 09/731,872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 459  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-876-997-459

Query Match 100.0%; Score 861; DB 10; Length 163;  
Best Local Similarity 100.0%; Pred. No. 9.7e-77;  
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60  
|  
Db 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60  
  
Qy 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120  
|  
Db 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120  
  
Qy 121 LSRDPLDNECGYQPPGAPGLGMPSSSCGPRRRRRDTRS 163  
|  
Db 121 LSRDPLDNECGYQPPGAPGLGMPSSSCGPRRRRRDTRS 163

## RESULT 3

US-09-978-360A-668  
; Sequence 668, Application US/09978360A  
; Publication No. US20040110939A1  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; APPLICANT: Clusel, Catherine  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
; FILE REFERENCE: 56 USA CIP  
; CURRENT APPLICATION NUMBER: US/09/978,360A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: US 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: US 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: US 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: US 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: US 60/099,273

; PRIOR FILING DATE: -09-04  
; PRIOR APPLICATION NUMBER: US 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: US 09/215,435  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: PCT/IB98/02122  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: US 09/247,155  
; PRIOR FILING DATE: 1999-02-09  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 810  
; SOFTWARE: Patent.pm  
; SEQ ID NO 668  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -52...-1  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (36, 47, 54, 65, 70, 103, 107, 108)  
; OTHER INFORMATION: unknown  
US-09-978-360A-668

Query Match 94.4%; Score 813; DB 11; Length 163;  
Best Local Similarity 94.5%; Pred. No. 5.3e-72;  
Matches 154; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
  
Qy 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60  
|  
Db 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60  
  
Qy 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120  
|  
Db 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120  
  
Qy 121 LSRDPLDNECGYQPPGAPGLGMPSSSCGPRRRRRDTRS 163  
|  
Db 121 LSRDPLDNECGYQPPGAPGLGMPSSSCGPRRRRRDTRS 163

## RESULT 4

US-10-425-115-237341  
; Sequence 237341, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 237341  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_148039C.1.pep  
US-10-425-115-237341

Query Match 35.4%; Score 305; DB 17; Length 72;  
Best Local Similarity 86.1%; Pred. No. 2.3e-22;  
Matches 62; Conservative 2; Mismatches 8; Indels 0; Gaps 0;  
  
Qy 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60  
|  
Db 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60







```

Db          40 MDRYQR-----VEKPRNDTPISQNEIRITTOGRMRNYISYGSMLLEENGHDEISIKAM 92
QY          61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVL 120
Db          93 GRAINKTMVVELIKRRVGGHQHTATESVDITDWEPL--EEGLLPLETRRHVSMTVT 150
QY          121 LSRDPLDPNCGYQPP 136
Db          151 LSKNPLDTSSPGYQPP 166

RESULT 14
US-10-425-114-60913
; Sequence 60913, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60913
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-105-A5_FLI.pep
US-10-425-114-60913

Query Match      23.6%; Score 203.5; DB 15; Length 325;
Best Local Similarity 33.8%; Pred. No. 1.5e-11;
Matches 46; Conservative 19; Mismatches 62; Indels 9; Gaps 2;

QY          1 MEHYRAGSVELPAPSPMPQPDPLEMRVRDGSKIRNLLGLGRLEGGSARHVVFSGS 60
Db          50 MDRYQR-----VEKPRNDTPISQNEIRITTOGRMRNYISYGSMLLEENGHDEISIKAM 102
QY          61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVL 120
Db          103 GRAINKTMVVELIKRRVGGHQHTATESVDITDWEPL--EEGLLPLETRRHVSMTVT 160
QY          121 LSRDPLDPNCGYQPP 136
Db          161 LSKNPLDTSSPGYQPP 176

RESULT 15
US-10-425-115-232929
; Sequence 232929, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 232929
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Zea mays

```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_144024C.1.pap
US-10-425-115-232929

Query Match      23.5%; Score 202; DB 17; Length 178;
Best Local Similarity 30.1%; Pred. No. 1e-11;
Matches 53; Conservative 23; Mismatches 70; Indels 30; Gaps 4;

Qy 1 MEHYKAGSVLPAPSPNPQLPPDTLEMRVVDGSKIRNLLGLALGRLEGGARHVFSGS 60
Db 1 MDRYQR-----VEKPRESPICANEIRITAOGRPRNYITYALALQDNATDDIVIKAM 53

Qy 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLQTESWVPASPDGTGLDPLTVRRHVPVAVVL 120
Db 54 GRAINKTVAVIVLLKRRRIAGLHNTSIESIDITDTWEPL--EEGLVTLETIRHVSLLITIK 111

Qy 121 LSRDPLDPNECGYOPP-----GAPFGLGMPSSSCGPRSRRRAR 159
Db 112 LSKKELDTSSFGYQPPIPADQVRPADFDQDAEAVPSGRG-----RGRGRSRGRGR 163

```

Search completed: January 26, 2005, 15:28:35  
Job time : 136.925 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 12:07:24 ; Search time 2520 Seconds  
(without alignments)  
9232.751 Million cell updates/sec

Title: US-10-057-813-13  
Perfect score: 492  
Sequence: 1 atggagcactaccggaagc.....ctcgagacacccgcgtga 492

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	492	100.0	871	9 BC032136	BC032136 Homo sapi
C 2	492	100.0	199114	9 AL1160270	AL1160270 Human DNA
3	484	98.4	884	6 BD139494	BD139494 Extended
4	374.4	76.1	461	6 AR412580	AR412580 Sequence
5	374.4	76.1	461	6 AX969414	AX969414 Sequence
6	374.4	76.1	461	6 BD108133	BD108133 EST and e
C 7	372	75.6	238811	2 AC098197	AC098197 Rattus no
C 8	372	75.6	242324	2 AC110351	AC110351 Rattus no
9	370.4	75.3	834	10 MM0293897	AJ293897 Mus muscu
10	370.4	75.3	863	10 BC047058	BC047058 Mus muscu
C 11	370.4	75.3	168268	10 AL807796	AL807796 Mouse DNA
12	236	48.0	303	6 AX887740	AX887740 Sequence
13	236	48.0	303	6 BD027350	BD027350 Sequence
14	236	48.0	305	6 BD077019	BD077019 5' EST of
15	85.4	17.4	1554	5 BC076268	BC076268 Danio rer
16	85.4	17.4	189318	2 CR376821	CR376821 Danio rer
C 17	85.4	17.4	213678	5 BX005012	BX005012 Zebrafish
18	81.6	16.6	242077	2 AC095805	AC095805 Rattus no
C 19	81.6	16.6	251145	2 AC108546	AC108546 Rattus no

20	77.8	15.8	1361	10	BC016085	BC016085 Mus muscu
21	74.6	15.2	981	6	CQ735345	CQ735345 Sequence
22	74.6	15.2	1183	9	AY034074	AY034074 Homo sapi
23	74.6	15.2	1473	9	BC007270	BC007270 Homo sapi
24	74.6	15.2	1483	9	BC002497	BC002497 Homo sapi
25	74.6	15.2	1941	6	AR338514	AR338514 Sequence
26	74.6	15.2	2375	6	AR447527	AR447527 Sequence
C 27	74.6	15.2	174559	9	AC015720	AC015720 Homo sapi
28	73	14.8	2015	9	AK000381	AK000381 Homo sapi
29	59.2	12.0	913	5	BX911336	BX911336 Gallus ga
30	55	11.2	125020	9	AF429315	AF429315 Homo sapi
C 31	55	11.2	167543	2	AC107974	AC107974 Homo sapi
32	53.2	10.8	1242	5	BX934615	BX934615 Gallus ga
C 33	48.2	9.8	65476	2	AC023362	AC023362 Mus muscu
34	46	9.3	799	3	BT011363	BT011363 Drosophil
C 35	45.4	9.2	110000	2	LMFLCHR34_13	Continuation (14 o
36	44.4	9.0	648	6	CQ592263	CQ592263 Sequence
37	44.4	9.0	2000	6	AX655393	AX655393 Sequence
C 38	43.4	8.8	125020	9	AF429315	AF429315 Homo sapi
39	42.8	8.7	882	3	AY089619	AY089619 Drosophil
C 40	41	8.3	2000	6	AX655393	AX655393 Sequence
C 41	40.6	8.3	182559	9	AC073476	AC073476 Homo sapi
C 42	40.6	8.3	231739	9	AC011811	AC011811 Homo sapi
C 43	40.4	8.2	244813	2	AC097868	AC097868 Rattus no
44	40.4	8.2	275726	2	AC118128	AC118128 Rattus no
C 45	40.4	8.2	300327	1	AE017228	AE017228 Mycobacte

## ALIGNMENTS

RESULT 1	BC032136	871 bp	mRNA	linear	PRI 30-JUN-2004
LOCUS	Homo sapiens chromosome 9 open reading frame 23, transcript variant 1, mRNA (CDNA clone MGC:29635 IMAGE:486926), complete cds.				
DEFINITION	BC032136				
ACCESSION	BC032136.2	GI:33879663			
VERSION	1	(bases 1 to 871)			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	(bases 1 to 871)			
AUTHORS	Klausner, R.D., Collins, F.S., Wagner, L.H., Grouse, L.H., Derge, J.G., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Shat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wille, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shetchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2	(bases 1 to 871)			
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 On Aug 19, 2003 this sequence version replaced gi:21618954.  
 Contact: MGC help desk  
 Email: [csapbs-remail.nih.gov](mailto:csapbs-remail.nih.gov)  
 Tissue Procurement: ATCC/DCTD/DTP  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McLoakey, J.C.,  
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
 Teurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 39 Row: h Column: 9  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 22325369.

FEATURES  
source

1..871  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:29635 IMAGE:4869926"  
 /tissue\_type="Skin, melanotic melanoma, high MDR."  
 /clone\_lib="NIH\_MGC\_49"  
 /lab\_host="DH10B-R"  
 /notes="Vector: pOTB7"

## gene

1..871  
 /gene="C9orf23"  
 /notes="synonyms: MGC29635, ba296122.5"  
 /db\_xref="LocusID:138716"  
 53..544

## CDS

/gene="C9orf23"  
 /codon\_start=1  
 /product="chromosome 9 open reading frame 23 protein"  
 /protein\_id="AAH32136.1"  
 /db\_xref="GI:21618955"  
 /db\_xref="LocusID:138716"  
 /translations="MEHYRKAGSVLPAPSPMPQLPDPDLEMRVDRGSKIRNLLGLAL  
 GRLEGSARHVFSGRAAGKAVCAEIVKRVPLGLHQLTKLRLFDSDVWPASPD  
 TGLDPLTVRRHVPAVWLLSRDPLDNECGYPPGAPPGLGMPSSSCGPRRRARD  
 TRS"

## ORIGIN

Query Match 100.0%; Score 492; DB 9; Length 871;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-114;  
 Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACACTACCGGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCAG 60  
 Db 53 ATGAGACACTACCGGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCAG 112  
 QY 61 CTACCTCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
 Db 113 CTACCTCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 172  
 QY 121 GGGTGGCTCTGGGTCTGGTGGGGCGGAGTGTCCGCATGTAGTGTCTCAGTTCT 180  
 Db 173 GGGTGGCTCTGGGTCTGGTGGGGCGGAGTGTCCGCATGTAGTGTCTCAGTTCT 232  
 QY 181 GGCAGGGCTGCAGGAAGGCTGTGACCTGCGCTGAGATTGTCAAGCGGGGTCCAGGC 240  
 Db 233 GGCAGGGCTGCAGGAAGGCTGTGACCTGCGCTGAGATTGTCAAGCGGGGTCCAGGC 292

QY 241 CTGACACAGCTACCAAGCTAGTTTCTTCTTCAAGTGAAGACAGCTGGGTCCAGGCTCA 300  
 Db 293 CTGACACAGCTACCAAGCTAGTTTCTTCTTCAAGTGAAGACAGCTGGGTCCAGGCTCA 352  
 QY 301 CTTGACACAGGGCTAGACCCCTCAGAGTGGCCCGCCATGTCTGCTGCTGCTGGTGGT 360  
 Db 353 CTTGACACAGGGCTAGACCCCTCAGAGTGGCCCGCCATGTCTGCTGCTGCTGGTGGT 412  
 QY 361 CTTGACCGGGACCCCTCGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420  
 Db 413 CTTGACCGGGACCCCTCGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 472  
 QY 421 GGCCTGGGTTCATGCCAGCTCCAGCTGTGCGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480  
 Db 473 GGCCTGGGTTCATGCCAGCTCCAGCTGTGCGCCCTCGTTCCCGAAGAGGGCTCGAGAC 532  
 QY 481 ACCCGATCGTGA 492  
 Db 533 ACCCGATCGTGA 544

RESULT 2  
AL160270/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

AL160270 199114 bp DNA linear PRI 27-APR-2001  
 Human DNA sequence from clone RP11-296L22 on chromosome 9, complete  
 sequence.  
 AL160270  
 AL160270.19 GI:13443330  
 HTG.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 199114)  
 Sehra, H.  
 Direct Submission  
 Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
 On Mar 24, 2001 this sequence version replaced gi:13274314.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em., EMBL; Sw.,  
 SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping  
 Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr9>  
 RP11-296L22 is from the library RPI-11.2 constructed by the group  
 of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-296L22 The true  
 left end of clone RP11-443p11 is at 176963 in this sequence. The  
 true right end of clone RP11-573M23 is at 15950 in this sequence.

## FEATURES

## source

1..199114  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"



```
repeat_region 45795..45925
/notes="AluSg repeat: matches 5..135 of consensus"
repeat_region 45953..46182
/notes="LINE2 repeat: matches 5933..6155 of consensus"

Query Match 100.0%; Score 492; DB 9; Length 199114;
Best Local Similarity 100.0%; Pred. No. 8.6e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
Db 195948 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 195889

QY 61 CTACCTCTGATACCTTGGATGCGGTCGAGATGGGAGCAAAATTCGCAACCTGCTG 120
Db 195888 CTACCTCTGATACCTTGGATGCGGTCGAGATGGGAGCAAAATTCGCAACCTGCTG 195829

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTGTCTCGGATGTAGTGTCTCAGGTTCT 180
Db 195828 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTGTCTCGGATGTAGTGTCTCAGGTTCT 195769

QY 181 GCGAGGGCTGCAGGAAGGCTGTACGTGGCTGAGATGTCAAGGCGGGTCCAGGC 240
Db 195768 GCGAGGGCTGCAGGAAGGCTGTACGTGGCTGAGATGTCAAGGCGGGTCCAGGC 195709

QY 241 CTGACACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGGAGCTGGTCCAGGCTCA 300
Db 195708 CTGACACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGGAGCTGGTCCAGGCTCA 195649

QY 301 CTGACACAGGCTAGACCCCTCACAGTGCAGGCGCCATGTGCTCAGTGTGGGTGCTG 360
Db 195648 CTGACACAGGCTAGACCCCTCACAGTGCAGGCGCCATGTGCTCAGTGTGGGTGCTG 195589

QY 361 CTCAGCGGAGCCCTTGAGCCCAATGATGTGTTACACCCCGAGGAGCAACCCCT 420
Db 195588 CTCAGCGGAGCCCTTGAGCCCAATGATGTGTTACACCCCGAGGAGCAACCCCT 195529

QY 421 GGCCTGGGTTCCATGCCAGCTCCAGTGTGGCCCTCGTTCCGAGAGAGGGCTCGAGAC 480
Db 195528 GGCCTGGGTTCCATGCCAGCTCCAGTGTGGCCCTCGTTCCGAGAGAGGGCTCGAGAC 195469

QY 481 ACCCGATCGTGA 492
Db 195468 ACCCGATCGTGA 195457

RESULT 3
BD139494
LOCUS
DEFINITION Extended cDNA of secretory protein.
ACCESSION BD139494
VERSION BD139494.1 GI:23234439
KEYWORDS JP 2002508182-A/246.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 884)
Bougueret,L., Duclert,A. and Edwards,J.B.D.M.
Extended cDNA of secretory protein
Patent: JP 2002508182-A 246 19-MAR-2002;
GENSET
OS Homo sapiens (human)
PN JP 2002508182-A/246
PD 19-MAR-2002
PF 17-DEC-1998 JP 2000539136
PR 17-DEC-1997 US 60/069957,09-FEB-1998 US 60/074121 PR
13-APR-1998 US 60/081563,10-AUG-1998 US 60/096116 PI
BOUGUERET,AYMERIC DUCLERT,JEAN BAPTISTE DUMAS MILNE PI EDWARDS
PC C12N15/09,C12N15/09,C07K14/47,C07K15/18,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/68,C12N15/00,C12N5/00,C12N15/00 CC
Von Heijne matrix
```

```
CC score 4.69999980926514
CC seq LGALGRLEGGSA/RH
FH Key Location/Qualifiers
FT CDS 69..557
FT sig_peptide 69..224
FT polyA_signal 849..854
FT polyA_site 870..883.
FT Location/Qualifiers
FEATURES
source
1..884
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 98.4%; Score 484; DB 6; Length 884;
Best Local Similarity 97.8%; Pred. No. 1.6e-112;
Matches 481; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
Db 69 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 128

QY 61 CTACCTCTGATACCTTTCAGATGCGGTCGAGATGGGAGCAAAATTCGCAACCTGCTG 120
Db 129 CTACCTCTGATACCTTTCAGATGCGGTCGAGATGGGAGCAAAATTCGCAACCTGCTG 188

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTGTCTCGGATGTAGTGTCTCAGGTTCT 180
Db 189 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTGTCTCGGATGTAGTGTCTCAGGTTCT 248

QY 181 GCGAGGGCTGCAGGAAGGCTGTACGTGGCTGAGATGTCAAGGCGGGTCCAGGC 240
Db 249 GCGAGGGCTGCAGGAAGGCTGTACGTGGCTGAGATGTCAAGGCGGGTCCAGGC 308

QY 241 CTGACACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGGAGCTGGTCCAGCCTCA 300
Db 309 CTGACACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGGAGCTGGTCCAGCCTCA 368

QY 301 CTGACACAGGCTAGACCCCTCACAGTGCAGGCGCCATGTGCTCAGTGTGGGTGCTG 360
Db 369 CTGACACAGGCTAGACCCCTCACAGTGCAGGCGCCATGTGCTCAGTGTGGGTGCTG 428

QY 361 CTCAGCGGAGCCCTTGAGCCCAATGATGTGTTACACCCCGAGGAGCAACCCCT 420
Db 429 CTCAGCGGAGCCCTTGAGCCCAATGATGTGTTACACCCCGAGGAGCAACCCCT 488

QY 421 GGCCTGGGTTCCATGCCAGCTCCAGTGTGGCCCTCGTTCCGAGAGAGGGCTCGAGAC 480
Db 489 GGCCTGGGTTCCATGCCAGCTCCAGTGTGGCCCTCGTTCCGAGAGAGGGCTCGAGAC 548

QY 481 ACCCGATCGTGA 492
Db 549 ACCCGATCGTGA 560

RESULT 4
AR412580
LOCUS
DEFINITION Sequence 217 from patent US 6639063.
ACCESSION AR412580
VERSION AR412580.1 GI:40167690
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 461)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 217 28-OCT-2003;
FEATURES
Location/Qualifiers
1..461
/organism="unknown"
/mol_type="genomic DNA"
```



## ORIGIN

Query Match 76.1%; Score 374.4; DB 6; Length 461;  
Best Local Similarity 98.4%; Pred. No. 1.3e-84;  
Matches 376; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGACACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGGCTTCCCAATGCCCCAG 60  
DB 78 ATGAGACACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGGCTTCCCAATGCCCCAG 137  
QY 61 CTACCTCTCTGATACCTTGGATCGGGTCCGAGATGGGAGTGGGCAAAATTCGCAACCTGCTG 120  
DB 138 CTACCTCTCTGATACCTTGGATCGGGTCCGAGATGGGAGTGGGCAAAATTCGCAACCTGCTG 197  
QY 121 GGGTTGGCTCTGGGTCGGTTCGAGGCGGCGAGTGTCTCGGATGTAGTGTCTCAGGTTCT 180  
DB 198 GGGTTGGCTCTGGGTCGGTTCGAGGCGGCGAGTGTCTCGGATGTAGTGTCTCAGGTTCT 257  
QY 181 GGCAGGCTGCGAGAAAGCTGTGAGTGGGTCGGATGGGAGTGGGCAAAATTCGCAACCTGCTG 240  
DB 258 GGCAGGCTGCGAGAAAGCTGTGAGTGGGTCGGATGGGAGTGGGCAAAATTCGCAACCTGCTG 317  
QY 241 CTGACACAGCTCACCAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 300  
DB 318 CTGACACAGCTCACCAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 377  
QY 301 CCTGACACAGGCTAGACCCCTTCACAGTGGCGGCCATGTGCTGCGAGTGTGGGTGCTG 360  
DB 378 CCTGACACAGGCTAGACCCCTTCACAGTGGCGGCCATGTGCTGCGAGTGTGGGTGCTG 437  
QY 361 CTCAGCGGAGACCCCTTGGACC 382  
DB 438 CTCAGCGGAGACCCCTTGGACC 459

## RESULT 6

BD108133  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BD108133  
EST and encoded human protein.  
BD108133  
BD108133.1 GI:23202951  
JP 2002010789-A/210  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 461)  
Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.  
EST and encoded human protein  
Patent: JP 2002010789-A 210 15-JAN-2002;  
GENSET CORP  
OS Homo sapiens (human)  
PN JP 2002010789-A/210  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 2000280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC Von Heijne matrix  
CC score 4.69999980926514  
CC seq LGLALRLEGGSA/RH  
FH Key Location/Qualifiers  
FT CDS 78..461  
FT sig\_peptide 78..233.  
FT Location/Qualifiers  
1..461  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## FEATURES

## source

1..461  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 76.1%; Score 374.4; DB 6; Length 461;  
Best Local Similarity 98.4%; Pred. No. 1.3e-84;  
Matches 376; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

## ORIGIN

Query Match 76.1%; Score 374.4; DB 6; Length 461;  
Best Local Similarity 98.4%; Pred. No. 1.3e-84;  
Matches 376; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

## sig\_peptide

78..233  
/notes="von Heijne matrix score 4.69999980926514 seq  
LGLALRLEGGSA/RH"

## CDS

1..461  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
78..461  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAFI4522.1"  
/db\_xref="GI:40975748"  
/translations="MEHYRKAGSVLPAPSPMPQLPDTLMRVRDGSKIRNLLGLAL  
GRLEGGARHVVFGSGRAAGKAVSCAEIVKRVPGHLQLTKLRFQTDTSWVPASPD  
TGLDPLTVRHVPVAVVLLSRXPWTQ"

## FEATURES

## Location/Qualifiers

1..461  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## REFERENCE

1  
Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.  
ESTs and encoded human proteins  
Patent: EP 1104808-A 217 06-JUN-2001;  
Genset (FR)

## AUTHORS

Homo sapiens (human)

## ORGANISM

Homo sapiens

## TITLE

Sequence 217 from Patent EP1104808.

## ACCESSION

AX969414

## VERSION

AX969414.1 GI:40975747

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## REFERENCE

1  
Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.  
ESTs and encoded human proteins  
Patent: EP 1104808-A 217 06-JUN-2001;  
Genset (FR)

## AUTHORS

Homo sapiens (human)

## ORGANISM

Homo sapiens

## TITLE

Sequence 217 from Patent EP1104808.

## ACCESSION

AX969414

## VERSION

AX969414.1 GI:40975747

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## REFERENCE

1  
Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.  
ESTs and encoded human proteins  
Patent: EP 1104808-A 217 06-JUN-2001;  
Genset (FR)

## AUTHORS

Homo sapiens (human)

## ORGANISM

Homo sapiens

## TITLE

Sequence 217 from Patent EP1104808.

## ACCESSION

AX969414

## VERSION

AX969414.1 GI:40975747

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## REFERENCE

1  
Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.  
ESTs and encoded human proteins  
Patent: EP 1104808-A 217 06-JUN-2001;  
Genset (FR)

## AUTHORS

Homo sapiens (human)

## ORGANISM

Homo sapiens

QY 1 ATGGAGCAGCTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCGCAG 60  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 78 ATGGAGCAGCTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCGCAG 137  
 QY 61 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 138 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 197  
 QY 121 GGGTGGCTGGCTGGTGGAGGGGGGAGTGGCTCGCATGTAGTGTCTCAGATTCT 180  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 198 GGGTGGCTGGCTGGTGGAGGGGGGAGTGGCTCGCATGTAGTGTCTCAGATTCT 257  
 QY 181 GGCAGGGCTGCAGAAAGCTGTGAGCTGGCTGAGATTGTCAAGCGGGGTCGCCAGGC 240  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 258 GGCAGGGCTGCAGAAAGCTGTGAGCTGGCTGAGATTGTCAAGCGGGGTCGCCAGGC 317  
 QY 241 GTGACAGCTACCAAGCTACCTTCTTCCAGCTGAGGACAGCTGGTCCGAGCTCA 300  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 318 CTGACAGCTACCAAGCTACCTTCTTCCAGCTGAGGACAGCTGGTCCGAGCTCA 377  
 QY 301 CCTGACAGGGCTAGACCCCTCAGCTGCGCGCCGATGTGCTCAGTGTGGTGGCTG 360  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 378 CCTGACAGGGCTAGACCCCTCAGCTGCGCGCCGATGTGCTCAGTGTGGTGGCTG 437  
 QY 361 CTGACGGGGACCCCTCGGACC 382  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 438 CTGACGGGGACCCCTCGGACC 459

## RESULT 7

AC098197/c  
 LOCUS AC098197  
 DEFINITION Rattus norvegicus clone CH230-9714, WORKING DRAFT SEQUENCE, 10  
 unordered pieces.  
 AC098197  
 AC098197.8 GI:30581495  
 HTG; HTGS PHAS1; HTGS DRAFT; HTGS\_FULLTOP.  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

## REFERENCE

AUTHORS Muzny, D. Marie, L., Metzger, M. Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Louseged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwono, G., Olarpunsgoon, A., Pal, S., Parks, K., Pasceriak, S., Paul, H., Perez, A., Perez, L., Ffankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 238811)  
 Worley, K. C.  
 Direct Submission  
 Submitted (23-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 238811)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On May 13, 2003 this sequence version replaced gi:24819301.  
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GIKF  
 Center clone name: CH230-9714  
 ----- Summary Statistics  
 Assembly program: Atlas 3.0;  
 Consensus quality: 228248 bases at least Q40  
 Consensus quality: 230539 bases at least Q30  
 Consensus quality: 230999 bases at least Q20  
 Estimated insert size: 236470; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 10 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 44755: contig of 44755 bp in length

\* 44756 44855: gap of unknown length  
\* 44856 129903: contig of 85048 bp in length  
\* 129904 130003: gap of unknown length  
\* 130004 227138: contig of 97135 bp in length  
\* 227139 227238: gap of unknown length  
\* 227239 228983: contig of 1745 bp in length  
\* 228984 230325: gap of unknown length  
\* 230326 230425: gap of unknown length  
\* 230426 232444: contig of 2019 bp in length  
\* 232445 232544: gap of unknown length  
\* 232545 234255: contig of 1711 bp in length  
\* 234256 234355: gap of unknown length  
\* 234356 235710: contig of 1355 bp in length  
\* 235711 235810: gap of unknown length  
\* 235811 236820: contig of 1010 bp in length  
\* 236821 236920: gap of unknown length  
\* 236921 238811: contig of 1891 bp in length.

## FEATURES

source Location/Qualifiers  
1. 238811  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-9714"  
misc\_feature 3538..4289  
/note="clone boundary  
clone\_end:Sp6  
site:EcORI  
end\_sequence:BH282839"  
misc\_feature 44856..46019  
/note="wgs\_contig"  
misc\_feature 128441..129903  
/note="wgs\_contig"

## ORIGIN

Query Match 75.68; Score 372; DB 2; Length 238811;  
Best Local Similarity 84.8; Pred. No. 2.9e-84;  
Matches 417; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 1 ATGAGCAGCTACCGAAAGCTGGCTGTAGAGCTCCAGCGCTCCCAATGCCCGAG 60  
Db 150054 ATGAGCAGCTACCGAGGGTGGCTGTAGAGCTCCAGCGCTATCACCAGTCCCCAG 149995  
QY 61 CTACCTCTGATACCTTGAGATCGGGTCCGAGATGCGAGCAAAATTCGCAACCTGCTG 120  
Db 149994 CTACCTCCGACACGCTGGAATCGCGTCCGAGATGGAGCAAAATCCGAACCTACTA 149935  
QY 121 GGGTGGCTGGTGGTGGAGGGCGGAGTGCTCGGATGTAGTGTCTCAGGTTCT 180  
Db 149934 GGGTGGCTGGTGGTGGAGGGAGGAGCAGCAGCGTGGTGTCTCAGGCTCC 149875  
QY 181 GGCAGGCTGCAGAAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGCGGGTCCGAGGC 240  
Db 149874 GGCAGGCTGCAGAAAGGCTGTGAGCTGCGAGATTGTCAAGCGCGGGTCCGAGGC 149815  
QY 241 CTCACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGCAGAGCTGGGTCCGAGCTCA 300  
Db 149814 CTCACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGCAGAGCTGGGTCCGAGCTCA 149755  
QY 301 CTTGACACAGGCTAGACCCCTTCAAGTGGCGGCGCATGTGCTGAGTGTGGGTGCTG 360  
Db 149754 CCGGACACGGGCTGGAGCCCTTCAAGTGGCGGCGCATGTGCTGAGTGTGGGTGCTG 149695  
QY 361 CTCAGCGGAGCCCTTGGAGCCCAATGAGTGTGGTGTACCAACCCCGAGGACCCCTT 420  
Db 149694 CTTAGTGGGACCCCTTGGAGCCCAATGAGTGTGGTGTACCAACCCCGAGGACCCCTT 149635  
QY 421 GGCCTGGGTTCATGCCAGCTCAGCTGTGGCTCTGTTCCGAGAGAGGGTCCGAGAC 480  
Db 149634 GGCCTGGGTTCATGCCAGCTCAGCTGTGGCTCTGTTCCGAGAGAGGGTCCGAGAC 149575  
QY 481 ACCGATCGTGA 492  
||||| |||

Db 149574 ACCCGTCTCTGA 149563

## RESULT 8

AC110351/c

LOCUS

DEFINITION

AC110351

AC110351.4 GI:23915435

HTG: HTGS PHASE2; HTGS DRAFT; HTGS\_FULLTOP.

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 242324)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M.,

Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorenshewa, L., Loubege, H., Lozano, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,

Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervil, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,

Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,

Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,

Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,

Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J.,

Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,

Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, Q.,

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,

Williams, G., Willson, R., Wlezyk, R., Wooden, H., Wortley, K.,

Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,

Yu, F., Zhang, J., Zhou, X., Zhou, S., Zhao, S., Dunn, D., von

Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,

Weinstock, G., and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 242324)

Worley, K. C.

Direct Submission

Submitted (11-FEB-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

JOURNAL

# REFERENCE AUTHORS TITLE JOURNAL

3 (bases 1 to 242324)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (12-OCT-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Oct 12, 2002 this sequence version replaced gi:21739791.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GRK  
Center clone name: CH230-217H18  
----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 222326 bases at least Q40  
Consensus quality: 224244 bases at least Q30  
Consensus quality: 225710 bases at least Q20  
Estimated insert size: 228121; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 242324: contig of 242324 bp in length.

## FEATURES

### source

Location/Qualifiers  
1..242324  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-217H18"

### misc\_feature

1..1109  
/note="wgs\_end\_extension  
clone end:T7"  
1611..3024  
/note="wgs\_end\_extension  
clone end:T7"

### misc\_feature

4338..42885  
/note="clone\_boundary  
clone end:T7  
site:EcoRI  
end\_sequence:RWBMD45TV"

### misc\_feature

45628..45664  
/note="clone\_boundary  
clone end:Sp6  
site:EcoRI  
end\_sequence:RWBMD45TV"

## ORIGIN

Query Match 75.6%; Score 372; DB 2; Length 242324;  
Best Local Similarity 84.8%; Pred. No. 2.9e-84;

Matches	417;	Conservative	0;	Mismatch	75;	Indels	0;	Gaps	0;
---------	------	--------------	----	----------	-----	--------	----	------	----

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCGAG 60  
 Db 13008 ATGGAGCACTACCGGAGGCTGGCTCTGTAGAGCTCCAGCGCCTCATCAATGCCCGAG 12949  
 QY 61 CTACCTCTGTATACCTTGTAGATGGGGTCCAGATGGAGCAAAATTCGCAACCTGTG 120  
 Db 12948 CTACCTCCGACACGCTGGAATGGGGTCCAGATGGAGCAAAATTCGCAACCTGTG 12889  
 QY 121 GGGTGGCTCTGGCTCGGTTGGAGGCGGCGAGTGTGCGCATGTAGTGTCTCAGTTCT 180  
 Db 12888 GGGCTGGCACTGGCTCGTGTGGAAGGAGGAGCAAGGACGTGGTGTCTCAGGCTCC 12829  
 QY 181 GGCAGGGCTGCAGGAAAGCTGTACGTCCGCTGTAGATTGTCAAGCGCGGGTCCAGGC 240  
 Db 12828 GGCAGGGCTGCAGGAAAGCGCTGAGTGTGAGAGATTGTCAAGCGCGGGTCCAGGC 12769  
 QY 241 CTGCACCAAGCTCACCAAGCTAGTTTCTTCAGACTGAGGACAGCTGGTCCAGCCTCA 300  
 Db 12768 CTACACCAAGCTCACCAAGCTTGGCTTCTGCAAACTGAGGACAGCTGGTCCCACTTCA 12709  
 QY 301 CCTGCACAGGGCTAGACCCCTCTACAGTGGCGGCCCATGTGCTGAGTGTGGTGTG 360  
 Db 12708 CCGGACAGGGCTGAGACCCCTCTACAGTCCGACGCCATGTGCTGAGTGTGGTGTG 12649  
 QY 361 CTACGCGGGACCCCTCTGAGCCCAATGAGTGTGGTGTACCAACCCCGAGGACACCCCT 420  
 Db 12648 CTTAGTGGGACCCCTCTGAGCCCAATGAGTGTGGTGTACCAACCCCGAGGACACCTCT 12589  
 QY 421 GGCCTGGGTTCATGCCAGCTCCAGCTGCGCCCTCGTTCGCGAGAGGGCTCGAGAC 480  
 Db 12588 GGCCTGGGTTCATGACCTAGTTCAGCTGTGCTCCAGACCCCGAGGAGGCTCGGAC 12529  
 QY 481 ACCCGATCGTGA 492  
 Db 12528 ACCCGTCTCTGA 12517

RESULT 9  
 MMU293897  
 LOCUS  
 DEFINITION  
 Mus musculus mRNA for hypothetical protein (ORF1), related to clone  
 Telethon(Italy\_B41)\_Strait02713\_F143.  
 ACCESSION  
 AJ293897  
 VERSION  
 AJ293897.1  
 KEYWORDS  
 ORF1.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1  
 Ievolella,C., Zara,I. and Lanfranchi,G.  
 Full length sequencing of some human and murine muscular transcript  
 (Telethon Italy project B41)  
 JOURNAL  
 Unpublished  
 REFERENCE  
 2 (bases 1 to 834)  
 Ievolella,C.  
 Direct Submission  
 TITLE  
 Submitted (01-SEP-2000) Ievolella C., CRIBI Biotechnology  
 Centre-Compl. Interdepartim. Vallisneri, Universita' di Padova, via  
 Ugo Bassi 58/b, 35121, ITALY  
 FEATURES  
 source  
 1..834  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /tissue\_type="skeletal muscle"  
 /tissue\_lib="Stratagene UNIZAP XR vector"  
 1..93  
 /note="ORF1"  
 94..585  
 /note="ORF1"

5'UTR  
 CDS

```

related to clone Telethon(Italy_B41)_Strait02713_Fl143"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC34588.1"
/db_xref="GI:113397925"
/translation="MEQYRRAGSVELPASPMPOLPDTLMEVRDGSKTRNLGLAL
GRLEGSTRAHVFGSGRAGKAVSCAEIYKRRVPLGLHQLTKLRPLQTESWVPTSPD
TGLDPLTVRRHVFVAVWVLLSRDPLDSECGYQPPAGPGIUGISIPSPSCGPRRRRRARD
TRS"
3'UTR 585..834
      /note="ORF1"
polyA_site 834
      /note="ORF1"

ORIGIN

Query Match      75.3%;   Score 370.4;   DB 10;   Length 834;
Best Local Similarity 84.6%;   Pred. No. 1.3e-83;
Matches 416;   Conservative 0;   Mismatches 76;   Indels 0;   Gaps 0;

QY 1 ATGAGGACATACCGGAAAGCTGGCTCTGTAGAGTCCACAGCGCTTCCCAATGCCCCAG 60
Db 94 ATGGAACAGTACCGGAGGGCGGCTCTGTAGAGCTTCACGCTCATCACAAATGCCCCAG 153
QY 61 CTACCTCTCATACCTTTGAGATGCGGGTCCGAGATGGCAGCAAAATCCGAACCTGCTG 120
Db 154 CTACCTCTGACACTCTGGAAATCGGAGTCCGAGATGGGAGCAAAATCCGAAACCTGCTA 213
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 214 GGGCTGGCGCTGGGTCTGTTGGAAAGGGGAAGCACAGGCACTGTGTGTCTCAGGCTCT 273
QY 181 GGCAGGGCTGCAGAAAGGTGTGACGTGCGCTGAGATTGTCAAGCGCGGGTCCCAAGGC 240
Db 274 GGCAGGGCTCTGGAAGAGGTGTGAGTGTGACAGATTGTCAAAAGCGGGTTCGGGGC 333
QY 241 CTCACAGCTACCAAGCTACGTTTTCCTTCAGACTGAGCAGCAGCTGGGTCCCAAGCTCA 300
Db 334 CTACACAGCTACCAAGCTCCGCTTCTCTGCAACGGAGGACAGCTGGGTCCCAAGCTCA 393
QY 301 CCTGCACACAGGGCTAGACCCCTCTCACAGTGGCGGCCCATGTGCTGCAGTGTGGGTGCTG 360
Db 394 CCAGACAGGGAATAGACCCCTCTCACAGTCCGAGCCACAGTGTGGGTGCTG 453
QY 361 CTAGCGGGGACCCCTGGAGCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCTCT 420
Db 454 CTTAGTCCGAGCCCTTGGAGCCCAAGTGTGGTACCAACCCCGAGGCTGACCTCTCT 513
QY 421 GGCTGGGTTCATGCCACCTCCAGCTGAGCTGGCCCTGTTTCCCGAAGAGGGCTCGAGAC 480
Db 514 GGCTGGGTTCATGACCTAGTCCAGCTGTGGCCCGGAGGAGGGGTTCGGGAC 573
QY 481 ACCCGATCGTGA 492
Db 574 ACCCGTCTGA 585

RESULT 10
BC047068
LOCUS
DEFINITION
IMAGE:5009034, complete cds.
ACCESSION
BC047068
VERSION
BC047068.1
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 863)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

```

## ORIGIN

Query Match	75.3%	Score 370.4	DB 10	Length 863
Best Local Similarity	84.6%	Pred. No. 1.3e-83		
Matches 416	Conservative 0	Mismatches 76	Indels 0	Gaps 0
QY	1	ATGGAGCAGTACCGGAAAGCTGGCTCTGTAGAGCTCCGAGCGCTTCCCAATGCCCCAG	60	
Db	91	ATGGAACAGTACCGGAGGCGGCTCTGTAGAGCTTCAGGCTCATCACCAATGCCCGAG	150	
QY	61	CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACTCGTGTG	120	
Db	151	CTACCTCCTGACACTCTGGAATCGAGTCCGAGATGGGACAAATCCGAAACTGCTA	210	
QY	121	GGGTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTTCCTCAGGTTCT	180	
Db	211	GGGTGGCGTGGTTCGTTTGGAAAGGGGAAGCACCGGACGCTGGTGTCTTCAGGCTCT	270	
QY	181	GGCAGGGCTCGAGGAAGGCTGTGAGTGGCTGGCTGAGATTGTCAAGCGCGGGTCCGAGGC	240	
Db	271	GGCGGGCTCTGGAAGGCTGTGAGTGGCTGGCTGAGATTGTCAAAAGCGCGGTCCGAGGC	330	
QY	241	CTGCACAGCTCACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCGAGCGCTCA	300	
Db	331	CTACACAGCTCACCAAGCTCGGTTTCTTGCAGGAGGACAGCTGGGTCCCAACTTCA	390	
QY	301	CCTGACACAGGGCTAGACCCCTTCAGTGGCGCCCATGTGCTCGAGTGTGGGTGTG	360	
Db	391	CCAGACAGGGACTTAGACCCCTTCACAGTCCGAGCCACGTCGCTGGGTGGGTACTG	450	
QY	361	CTCAGCGGGAGCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT	420	
Db	451	CTTAGTGGGACCCCTTGGACCCCGAGTGAATGTGGCTACCAACCCCGAGGTGCACCTCT	510	
QY	421	GGCTTGGTTCCATGCCAGCTCCAGTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC	480	
Db	511	GGCTTGGTTCCATACCTAGTCCAGCTGTGGCCCCAGACCCCGAAGAGGGCTCGGAC	570	
QY	481	ACCCGATCGTGA	492	
Db	571	ACCCGGTCTTGA	582	

**RESULT** 11  
AL807796/c  
**LOCUS**

**DEFINITION** Mouse DNA sequence from clone RP23-167I12 on chromosome 4, complete sequence.  
Al807796            168268 bp         DNA          linear      ROD 04-FEB-2003

**ACCESSION**  
Mouse DNA sequence from clone RP23-167I12 on chromosome 4, complete sequence.

**VERSION**  
HTG.

**KEYWORDS**  
Mus musculus (house mouse)

**SOURCE**  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE**  
1 (bases 1 to 168268)  
Harrison, E.

**AUTHORS**  
Direct Submission

**TITLE**  
Submitted (04-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerv@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

**JOURNAL**  
On Jan 30, 2003 this sequence version replaced gi:27653683.  
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been seen where it has a phred quality of at least 30.

**COMMENT**  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquerv@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/Celegans/wormpep> RP23-167112 is from the RPC1-23 Mouse BAC Library constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

```
Location/Qualifiers
1. .168268
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="4"
   /clone="RP23-167112"
   /clone_lib="RP23-167112"
```

## ORIGIN

Query Match	75.3%;	Score	370.4;	DB	10;	Length	168268;
Best Local Similarity	84.6%;	Pred. No.	7.8e-84;				
Matches	416;	Conservative	0;	Mismatches	76;	Indels	0;
QY	1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGAGCCCTTCCCAATGCCCCAG	60				
Db	69359	ATGGAACAGTACCGGAGGCCGGCTCTGTAGAGCTTCCAGCTCATCACCAATGCCCCAG	693000				
QY	61	CTACCTCTCTGAYACCCCTTGAGATGGGGTCCGAGATGGCAGCAAAATTCCGCAACCTGCTG	120				
Db	69299	CTACCTCTCTGACACTCTGGAATGCGAGTCCGAGATGGGAGCAAAATCCGAAACCTGCTA	69240				
QY	121	GGGTGGCTCTGGTCCGTTGAGAGGGCGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT	180				
Db	69239	GGGCTGGCCCTGGTCTGTTTGAAGGGGGAAGCACAGGACAGTGTGTCTCAGGCTCT	69180				
QY	181	GGCAGGGCTGCAGGAAAGGCTCTCAGCTGCGCTGAGATGTCAAGGGGGGTCCAGGC	240				
Db	69179	GGCCGGGCTGCTGGAAAGGCTGTCACTGTGCAGAGATGTCAAAGGGGGGTTCGGGC	69120				
QY	241	CTGCACAGCTCACCAAGTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCCGAGCTCA	300				
Db	69119	CTACACCAAGCTCACCAAGCTCCGCTTCTGTGCAAACGAGGACAGCTGGTCCCAACTTCA	69060				
QY	301	CTGTGACACAGGGCTAGACCCCTCTCAAGTGCOCGCCCATGTGCCCTGCAGTGTGGGTGTG	360				
Db	69059	CCAGACACGGGACTAGACCCCTCTCAAGTCCGACCGACCTGCCGTGGGTGTGGGTACTG	69000				
QY	361	CTCAGCCGGGACCCCTTGAGCCCAATGAGTGTGGTTACCAACCCCCAGGAGCACCCCT	420				
Db	68999	CTTAGTCGGGACCCCTTGAGCCCAAGTGAATGTGGCTACCAACCCCCAGGTCATCTCT	68940				
QY	421	GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTTCCGGAAGAGGGCTCGAGAC	480				
Db	68939	GGCCTGGGTTCCATACCTAGTCCCAAGTGTGGCCCCAGACCCCGAAGGAGGGCTCGGAC	68880				
QY	481	ACCGATCGTGA	492				
Db	68879	ACCCGGTCTGA	68868				

```
RESULT 12
AX887740          303 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION       Sequence 3603 from Patent EP1033401.
ACCESSION        AX887740
VERSION          AX887740.1 GI:40046203
KEYWORDS
SOURCE           Homo sapiens (human)
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS          Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE            Expressed sequence tags and encoded human proteins
JOURNAL          Patent: EP 1033401-A 3603 06-SEP-2000;
                  Genset (FR)
FEATURES         Location/Qualifiers
source           1..303
                 /organism="Homo sapiens"
                 /mol_type="unassigned DNA"
                 /db_xref="taxon:9606"
                 68..>301
                 /note="unnamed protein product"
                 /codon_start=1
                 /protein_id="CAFO1878.1"
                 /db_xref="GI:40046204"
                 /translations="MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLAL
ORIGIN
Query Match      48.0%; Score 236; DB 6; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.3e-49;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCCAG 60
Db 68 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCCAG 127

QY 61 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 128 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 187

QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGCATGTAGTGTCTCAGGTTCT 180
Db 188 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGCATGTAGTGTCTCAGGTTCT 247

QY 181 GGCAGGCTCGAGAAAGCTGTGAGTCTCGCTGAGATTGTCAAGCGCGGGTCCC 236
Db 248 GGCAGGCTCGAGAAAGCTGTGAGTCTCGCTGAGATTGTCAAGCGCGGGTCCC 303

RESULT 13
BD027350          303 bp      DNA      linear      PAT 27-AUG-2002
LOCUS            BD027350
DEFINITION       Sequence tag and encoded human protein.
ACCESSION        BD027350
VERSION          BD027350.1 GI:22569092
KEYWORDS         JP 2001269182-A/3596.
SOURCE           Homo sapiens (human)
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS          Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE            Sequence tag and encoded human protein
JOURNAL          Patent: JP 2001269182-A 3596 02-OCT-2001;
                  GENSET
COMMENT          OS Homo sapiens (human)
                  PN JP 2001269182-A/3596
                  PD 02-OCT-2001
                  PF 24-FEB-2000 JP 2000118773
                  PR 26-FEB-1999 US 60/122487
                  PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES

PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers
FT CDS          1..303
                 /organism="Homo sapiens"
                 /mol_type="genomic DNA"
                 /db_xref="taxon:9606"
FEATURES         source
ORIGIN
Query Match      48.0%; Score 236; DB 6; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.3e-49;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCCAG 60
Db 68 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCCAG 127

QY 61 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 128 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 187

QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGCATGTAGTGTCTCAGGTTCT 180
Db 188 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGCATGTAGTGTCTCAGGTTCT 247

QY 181 GGCAGGCTCGAGAAAGCTGTGAGTCTCGCTGAGATTGTCAAGCGCGGGTCCC 236
Db 248 GGCAGGCTCGAGAAAGCTGTGAGTCTCGCTGAGATTGTCAAGCGCGGGTCCC 303

RESULT 14
BD077019          305 bp      DNA      linear      PAT 27-AUG-2002
LOCUS            BD077019
DEFINITION       5' EST of secretory protein expressed in prostate.
ACCESSION        BD077019
VERSION          BD077019.1 GI:22622622
KEYWORDS         JP 2001512013-A/266.
SOURCE           Homo sapiens (human)
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS          Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
TITLE            5' EST of secretory protein expressed in prostate
JOURNAL          Patent: JP 2001512013-A 266 21-AUG-2001;
                  GENSET
COMMENT          OS Homo sapiens (human)
                  PN JP 2001512013-A/266
                  PD 21-AUG-2001
                  PF 31-JUL-1998 JP 2000505291
                  PR 01-AUG-1997 US 08/905144
                  PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO PI
                  LACROIX
                  PC C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
                  blasn
                  CC identity 99
                  CC region 9..294
                  CC id W52125
                  CC est
                  CC blastn
                  CC identity 100
                  CC region 1..265
                  CC id AA024623
                  CC est
                  CC blastn
                  CC identity 100
                  CC region 1..263
```

```

CC id H55824
CC est
CC blastn
CC identity 97
CC region 2. .288
CC id R62921
CC est
CC blastn
CC identity 99
CC region 57. .242
CC id N31702
CC est
CC blastn
CC identity 100
CC region 1. .56
CC id N31702
CC est
CC Von Heijne matrix
CC score 4.7
CC seq LGALGRLEGGSA/RH
FH Key Location/Qualifiers
FT misc_feature 2. .287
FT misc_feature 19. .283
FT misc_feature 22. .284
FT misc_feature 21. .307
FT misc_feature 102. .287
FT misc_feature 45. .100
FT sig_peptide 69. .224.
FEATURES
    source
        1. .305
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match 48.0%; Score 236; DB 6; Length 305;
Best Local Similarity 100.0%; Pred. No. 2.3e-49;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCACTACCGGAAGCTGCTCTGTAGACTCCAGCGCTTCCCAATGCCCGAG 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 69 ATGAGCACTACCGGAAGCTGCTCTGTAGACTCCAGCGCTTCCCAATGCCCGAG 128
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 CTACCTCTCATACCTTGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 129 CTACCTCTCATACCTTGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 188
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 GGGTTGGCTCGGTTGGAGGGCGGAGTCTCGGCATGCTAGTGTCTCAGTTCTT 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 189 GGGTTGGCTCGGTTGGAGGGCGGAGTCTCGGCATGCTAGTGTCTCAGTTCTT 248
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 GGCAGGCTCAGGAAGGCTGTGAGTGGCTGAGATGTTCAAGCGGGGTCCC 236
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 249 GGCAGGCTCAGGAAGGCTGTGAGTGGCTGAGATGTTCAAGCGGGGTCCC 304
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
BC076268 1554 bp mRNA linear VRT 27-JUL-2004
LOCUS Danio rerio zgc:92794, mRNA (cdna clone MGC:92794 IMAGE:7088292),
DEFINITION complete cds.
ACCESSION BC076268
VERSION BC076268.1 GI:49901115
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
    Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
    Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
    Cypriniformes; Cyprinidae; Danio.
    1 (bases 1 to 1554)
    Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
    Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
    Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

```

## ORIGIN

Query Match

17.4%; Score 85.4; DB 5; Length 1554;

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toehiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 1554)  
Strausberg, R.  
Direct Submission  
Submitted (06-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: John Ngai, Univ of CA, Berkeley  
cDNA Library Preparation: Dr. Sumio Sugano  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 180 Row: j Column: 2  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

## FEATURES

source

```

1. .1554
    /organism="Danio rerio"
    /mol_type="mRNA"
    /db_xref="taxon:7955"
    /clone="MGC:92794 IMAGE:7088292"
    /tissue_type="Brain"
    /clone_lib="NIH_ZGC_4"
    /lab_host="DH10B"
    /note="Vector: pME18S-FL3"
1. .1554
    /gene="zgc:92794"
    /db_xref="LocusID:436813"
243. .917
    /gene="zgc:92794"
    /codon_start=1
    /product="zgc:92794"
    /protein_id="AAH76268.1"
    /db_xref="GI:49901115"
    /db_xref="LocusID:436813"
    /translation="MENYRKANTIEQPCPCFPDLPSTPEVRVKDGSKIRNLMPFAL
SRMEETAAADHEGSVGVGTGDLNLCVQIVFTGVGQSVAKAITCEVIMKRRIHGLHQ
LTKLAYRTLQDWVEPLEPGALDLSVSRNVPSIWLVSRLSDLSQPCYQAPGSGFDA
LWIOALKEEAATPRHGQRKRGGTGTGAEVGGRGKPRKPRGPDTRKPPGHAGGG
QEGMNV"

```

## gene

## CDS



Best Local Similarity 58.4%; Pred. No. 4.1e-11;  
Matches 149; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy	156	TCGGCATGTAGTGTCTCAGGTTCTGGCAGGGCTGCAGGAAAGGCTGTCTAGCTGCGCTGA	215
Db	455	TCGCCAATAGTATTTCAGGCGTGGGTCAAGCGTTGCAAAAGCAATCAGTGTGTAGA	514
Qy	216	GATTGTCAAGCGGCGGTCCAGGCTGCCAGGCTGCACCAAGCTACGTTTCCTTCAGAC	275
Db	515	GATCATGAACGGCGCATTCATGGCTACATCAGCTCACTAAGCTGGCCTACCGCACACT	574
Qy	276	TGAGGACAGTGGTCCAGGCTCACCTGACACAGGGCTAGACCCCTCAGTGGCCCG	335
Db	575	TCAGGATGTCTGGAGAGCCGCTGGAGCCTGGGGCGGGCTGGACAGTCTCACCGTCAGCAG	634
Qy	336	CCATGTCCCTGCAGTGTGGGTGTCTCAGCGGGACCCCTTGACCCCAATGAGTGTGG	395
Db	635	AAATGTACCTAGTATTGGGTTCTTCTCTCAAGAGACTCGCTTGTATGAACCAAGCCGGG	694
Qy	396	TTACCAACCCCCAGG	410
Db	695	TTATCAAGCTCCAGG	709

Search completed: January 31, 2005, 17:27:11  
Job time : 2526 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 12:10:30 ; Search time 2432 Seconds  
(without alignments)  
7371.854 Million cell updates/sec

Title: US-10-057-813-13

Perfect score: 492  
Sequence: 1 atggagcactaccggaagc.....ctcgagacacccgatcgta 492

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_ges1.\*  
9: gb\_ges2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	492	100.0	562	5	BU785328 in44b11.y
2	492	100.0	631	2	AW957840 EST369910
3	492	100.0	711	4	BG328094 602427145
4	492	100.0	791	6	CB959493 AGENCOURT
5	492	100.0	888	1	AL528391
6	492	100.0	888	3	CR619309 full-length
7	492	100.0	918	5	BQ218097 AGENCOURT
8	492	100.0	940	5	BQ643214 AGENCOURT
9	492	100.0	1000	5	BQ056655 AGENCOURT
10	492	100.0	1025	5	BQ056288 AGENCOURT
11	492	100.0	1074	2	BE742621
12	492	100.0	1079	5	BQ063771
13	492	100.0	1104	4	BM806218 AGENCOURT
14	491	99.8	967	5	BQ070451 AGENCOURT
15	490.4	99.7	604	6	CB110489 K-EST0151
16	490.4	99.7	964	4	BG766764
17	490.4	99.7	1171	4	BM467200 AGENCOURT
18	490	99.6	734	1	AV705461
19	489.4	99.5	572	7	CK902681 in44b11.x
20	486.2	98.8	889	5	BUL151215 AGENCOURT
21	485.6	98.7	913	5	BQ887804 AGENCOURT
22	483	98.2	717	4	BM701721 UI-E-CQ1
23	479.4	97.4	830	4	BG766574 602739343
24	479	97.4	633	2	AW957916 EST369986

25	473	96.1	566	7	CN406927	CN406927	170006000
26	469.8	95.5	526	1	AA313780	EST185643	
27	468	95.1	492	5	BU076706	ims1e03.y	
28	462	93.9	530	1	AA469996	zu10c07.x	
29	461.6	93.8	948	4	BI768478	603053963	
30	459	93.3	923	2	BE790325	601480501	
31	458.4	93.2	978	5	BU159046	AGENCOURT	
32	445	90.4	890	2	BE899574	601682407	
33	438.4	89.1	780	2	BF699406	602125986	
34	437.8	89.0	887	5	BQ224958	AGENCOURT	
35	434.8	88.4	790	2	BE897665	601438569	
36	426	86.6	494	5	EX280954	EX280954	
37	410.4	83.4	741	5	BQ575028	UI-H-EZ1-	
38	408.8	83.1	857	7	CN157400	946103 MA	
39	408.8	83.1	870	7	CN159345	948407 MA	
40	407.2	82.8	573	4	BG733063	346756 MA	
41	405.6	82.4	662	5	BP453103	BP453103	
42	405.6	82.4	740	5	BU689983	UI-CF-PNO	
43	405.6	82.4	774	5	CK454492	914978 MA	
44	405.6	82.4	774	5	BX671114	BX671114	
45	403.8	82.1	580	1	AA206557	zq51d08.x	

#### ALIGNMENTS

RESULT 1  
BU785328  
LOCUS  
DEFINITION in44b11.y1 HR85 islet Homo sapiens mRNA linear EST 11-OCT-2002  
similar to TR:Q9V9B6 Q9V9B6 CG9422 PROTEIN. 1, mRNA sequence.  
ACCESSION BU785328  
VERSION BU785328.1 GI:23831081  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 562)  
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T., Jackson,Y. and Bowers,Y.  
TITLE Endocrine Pancreas Consortium  
JOURNAL Unpublished (2000)  
COMMENT Other\_ESTs: in44b11.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohpc.harvard.edu  
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 445.

#### FEATURES

Location/Qualifiers  
1..562  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6124773"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_lib="HR85 islet"  
/notes="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1: NotI; Site\_2: XhoI; cDNA made by oligo-dT priming."

Size-selected on agarose gel. Average insert size ~1kb. 5' xhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

## ORIGIN

Query Match 100.0%; Score 492; DB 5; Length 562;  
Best Local Similarity 100.0%; Pred. No. 2.7e-115;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGACCTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 9 ATGAGGACCTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 68

QY 61 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 69 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 128

QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
Db 129 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 188

QY 181 GGCAGGGCTCGAGGAAGGCTGTGAGCTCGCTGAGATTGTCAAGCGCGGGTCCAGGC 240  
Db 189 GGCAGGGCTCGAGGAAGGCTGTGAGCTCGCTGAGATTGTCAAGCGCGGGTCCAGGC 248

QY 241 CTGACACAGCTACCAAGTACGTTTCTTTTTCAGCTGAGGACAGCTGGGTCCAGGCTCA 300  
Db 249 CTGACACAGCTACCAAGTACGTTTCTTTTTCAGCTGAGGACAGCTGGGTCCAGGCTCA 308

QY 301 CCTGACACAGGCTAGACCCCTTACAGTGGCGGCGCATGTGCTGAGTGGGTGCTG 360  
Db 309 CCTGACACAGGCTAGACCCCTTACAGTGGCGGCGCATGTGCTGAGTGGGTGCTG 368

QY 361 CTGACCGGACCCCTTGGACCCCAATGAGTGTGTTTACCAACCCCGAGGACACCCCT 420  
Db 369 CTGACCGGACCCCTTGGACCCCAATGAGTGTGTTTACCAACCCCGAGGACACCCCT 428

QY 421 GGCCTGGTTCATGCCCAGCTCCAGCTGGGCGCTGTTTCCGAGAGAGGCTCGAGAC 480  
Db 429 GGCCTGGTTCATGCCCAGCTCCAGCTGGGCGCTGTTTCCGAGAGAGGCTCGAGAC 488

QY 481 ACCCGATCGTGA 492  
Db 489 ACCCGATCGTGA 500

RESULT 2  
AW957840 631 bp mRNA linear EST 01-JUN-2000  
LOCUS  
DEFINITION  
EST369910 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
AW957840  
VERSION  
AW957840.1 GI:8147523  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 631)  
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and Quackenbush, J.  
Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray  
Unpublished (2000)  
Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208

Email: johnq@tigr.org  
Plate: 111

Seq primer: Reverse.  
Location/Qualifiers

FEATURES  
source

1. .631  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGE"  
/note="Vector: pBluescriptSKm"

## ORIGIN

Query Match 100.0%; Score 492; DB 2; Length 631;  
Best Local Similarity 100.0%; Pred. No. 2.7e-115;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGACCTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 52 ATGAGGACCTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 111

QY 61 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 112 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 171

QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
Db 172 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 231

QY 181 GGCAGGGCTCGAGGAAGGCTGTGAGCTCGCTGAGATTGTCAAGCGCGGGTCCAGGC 240  
Db 232 GGCAGGGCTCGAGGAAGGCTGTGAGCTCGCTGAGATTGTCAAGCGCGGGTCCAGGC 291

QY 241 CTGACACAGCTACCAAGTACGTTTCTTTTTCAGCTGAGGACAGCTGGGTCCAGGCTCA 300  
Db 292 CTGACACAGCTACCAAGTACGTTTCTTTTTCAGCTGAGGACAGCTGGGTCCAGGCTCA 351

QY 301 CCTGACACAGGCTAGACCCCTTACAGTGGCGGCGCATGTGCTGAGTGGGTGCTG 360  
Db 352 CCTGACACAGGCTAGACCCCTTACAGTGGCGGCGCATGTGCTGAGTGGGTGCTG 411

QY 361 CTGACCGGACCCCTTGGACCCCAATGAGTGTGTTTACCAACCCCGAGGACACCCCT 420  
Db 412 CTGACCGGACCCCTTGGACCCCAATGAGTGTGTTTACCAACCCCGAGGACACCCCT 471

QY 421 GGCCTGGTTCATGCCCAGCTCCAGCTGGGCGCTGTTTCCGAGAGAGGCTCGAGAC 480  
Db 472 GGCCTGGTTCATGCCCAGCTCCAGCTGGGCGCTGTTTCCGAGAGAGGCTCGAGAC 531

QY 481 ACCCGATCGTGA 492  
Db 532 ACCCGATCGTGA 543

RESULT 3  
BG328094 711 bp mRNA linear EST 27-FEB-2001  
LOCUS  
DEFINITION  
602427145F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4546643 5', mRNA sequence.  
ACCESSION  
BG328094  
VERSION  
BG328094.1 GI:13134532  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 711)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D..  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LNCM233 row: i column: 12  
High quality sequence stop: 706.

## FEATURES

source  
1. .711  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4546643"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 15"  
/notes="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; CDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)"

## ORIGIN

Query Match 100.0%; Score 492; DB 4; Length 711;  
Best Local Similarity 100.0%; Pred. No. 2.7e-115;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCACGGCTTCCCAATGCCCCAG 60  
DB 54 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCACGGCTTCCCAATGCCCCAG 113  
QY 61 CTACCTCTCATACCTTGGATGCGGTCGGATGGGAGGAGCAAAATTCGACACCTGCTG 120  
DB 114 CTACCTCTCATACCTTGGATGCGGTCGGATGGGAGGAGCAAAATTCGACACCTGCTG 173  
QY 121 GGGTTGGCTCTGGGTGCGTTGGAGGGCGGCGAGTGCTCGGCATGTAGTGTCTCAGGTTCT 180  
DB 174 GGGTTGGCTCTGGGTGCGTTGGAGGGCGGCGAGTGCTCGGCATGTAGTGTCTCAGGTTCT 233  
QY 181 GGACGGCTCGAGGAAAGGCTGTACGTGCGCTGAGATGTCAAGCGGGGTCGCCAGGC 240  
DB 234 GGACGGCTCGAGGAAAGGCTGTACGTGCGCTGAGATGTCAAGCGGGGTCGCCAGGC 293  
QY 241 CTCACACAGCTACCAAGCTACGTTTCTTCAGACTGAGACAGAGCTGGGTCCAGCCTCA 300  
DB 294 CTCACACAGCTACCAAGCTACGTTTCTTCAGACTGAGACAGAGCTGGGTCCAGCCTCA 353  
QY 301 CCTGACACAGGCTAGACCCCTCTCACAGTGGCGCGCATGTGCTGCGATGTGGGTGCTG 360  
DB 354 CCTGACACAGGCTAGACCCCTCTCACAGTGGCGCGCATGTGCTGCGATGTGGGTGCTG 413  
QY 361 CTCAGCGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGAGACACCCCT 420  
DB 414 CTCAGCGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGAGACACCCCT 473  
QY 421 GGCTGGGTTCCATGCCAGCTCCAGTGTGGCCCTCGTTCCGAGAGAGGGCTCAGAC 480  
DB 474 GGCTGGGTTCCATGCCAGCTCCAGTGTGGCCCTCGTTCCGAGAGAGGGCTCAGAC 533  
QY 481 ACCGATCGTGA 492  
DB 534 ACCGATCGTGA 545

## RESULT 4

CB959493  
LOCUS  
DEFINITION  
CB959493 791 bp mRNA linear EST 29-APR-2003  
AGENCOURT\_13887473 NIH\_MGC\_147 Homo sapiens cDNA clone  
IMAGE:30341789 5', mRNA sequence.  
ACCESSION  
CB959493

## VERSION

KEYWORDS  
SOURCE  
ORGANISM

CB959493.1 GI:30215609

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 791)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Dr. Stefan Hansson

CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help

and advice from Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NDAM373 row: n column: 06

High quality sequence stop: 631.

Location/Qualifiers

1. .791

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30341789"

/tissue\_type="Human Placenta"

/lab\_host="DH10B TON"

/clone\_lib="NIH MGC 147"

/note="Organ: Placenta; Vector: pBluescriptR; Site 1:

all-XhoI; Site 2: BamH; oligo-dT primed using primer

5'-TTTTTTTTTTTNN-3', size-selected for average

insert size 2.3 kb and normalized to 10<sup>6</sup> copies

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: This is

a NIH\_MGC library."

ORIGIN

Query Match 100.0%; Score 492; DB 6; Length 791;  
Best Local Similarity 100.0%; Pred. No. 2.8e-115;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCACGGCTTCCCAATGCCCCAG 60  
DB 95 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCACGGCTTCCCAATGCCCCAG 154  
QY 61 CTACCTCTGTATACCTTCAGATCGGGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 120  
DB 155 CTACCTCTGTATACCTTCAGATCGGGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 214  
QY 121 GGGTTGGCTCTGGGTGCGTTGGAGGGCGGCGAGTGCTCGGCATGTAGTGTCTCAGGTTCT 180  
DB 215 GGGTTGGCTCTGGGTGCGTTGGAGGGCGGCGAGTGCTCGGCATGTAGTGTCTCAGGTTCT 274  
QY 181 GGACGGCTCGAGGAAAGGCTGTACGTGCGTGTAGATGTGTCAAGCGGGGTCGCCAGGC 240  
DB 275 GGACGGCTCGAGGAAAGGCTGTACGTGCGTGTAGATGTGTCAAGCGGGGTCGCCAGGC 334  
QY 241 CTGACACAGCTACCAAGCTAGTGTTCCTTCAGACTGAGAGACAGCTGGGTCCAGCCTCA 300  
DB 335 CTGACACAGCTACCAAGCTAGTGTTCCTTCAGACTGAGAGACAGCTGGGTCCAGCCTCA 394  
QY 301 CCTGACACAGGCTAGACCCCTCTCACAGTGGCGCGCATGTGCTCAGTGTGGGTGCTG 360  
DB 395 CCTGACACAGGCTAGACCCCTCTCACAGTGGCGCGCATGTGCTCAGTGTGGGTGCTG 454  
QY 361 CTCAGCGGACCCCTGGACCCCAATGAGTGTGTTACCAACCCCGAGAGACACCCCT 420  
DB 455 CTCAGCGGACCCCTGGACCCCAATGAGTGTGTTACCAACCCCGAGAGACACCCCT 514

```

QY 421 GGCCTGGTTCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGCTCGAGAC 480
    |||
Db 515 GGCCTGGTTCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGCTCGAGAC 574

QY 481 ACCCGATCGTGA 492
    |||
Db 575 ACCCGATCGTGA 586

RESULT 5
AL528391
LOCUS
DEFINITION
AL528391 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC029YK21 5-PRIME, mRNA sequence.
ACCESSION
AL528391
VERSION
AL528391.3 GI:45703459
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 888)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:31066241.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5101.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DC029AF11QP1&c=5101.r.
FEATURES
    source
    1..888
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0DC029YK21"
        /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
        /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT)
        primer. Five prime end enriched, double-strand cDNA was
        digested with Not I and EcoR V sites of the Not I and EcoR V
        sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 100.0%; Score 492; DB 1; Length 888;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
    |||
Db 87 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 146

QY 61 CTACCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
    |||
Db 147 CTACCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 206

QY 121 GGGTGGCTCTGGGTGGGTGGGCGGCGGAGTGTCCGCATGTAGTGTCTCAGGTTCT 180
    |||
Db 207 GGGTGGCTCTGGGTGGGTGGGCGGCGGAGTGTCCGCATGTAGTGTCTCAGGTTCT 266

QY 181 GGCAGGGCTGCAGGAAGGCTGTACGCTCGCTGAGATTGTCAAGCGGGGGTCCAGGC 240
    |||
Db 267 GGCAGGGCTGCAGGAAGGCTGTACGCTCGCTGAGATTGTCAAGCGGGGGTCCAGGC 326

```

```

QY 241 CTGCACCACTCACCAGCTACGTTTCCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 300
    |||
Db 327 CTGCACCACTCACCAGCTACGTTTCCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 386

QY 301 CTGTACACAGGGGTAGACCCCTCTACAGTGCCTCCGCGCATGTGCTGCTGGGTGGTGG 360
    |||
Db 387 CTGTACACAGGGGTAGACCCCTCTACAGTGCCTCCGCGCATGTGCTGCTGGGTGGTGG 446

QY 361 CTGAGCCGGGACCCCTCTGAGACCCCAATGATGTGTTACCAACCCCGAGGAGCACCCCT 420
    |||
Db 447 CTGAGCCGGGACCCCTCTGAGACCCCAATGATGTGTTACCAACCCCGAGGAGCACCCCT 506

QY 421 GGCCTGGGTTCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGCTCGAGAC 480
    |||
Db 507 GGCCTGGGTTCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGCTCGAGAC 566

QY 481 ACCCGATCGTGA 492
    |||
Db 567 ACCCGATCGTGA 578

RESULT 6
CR619309
LOCUS
DEFINITION
CR619309 full-length cDNA clone CS0DC029YK21 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).
ACCESSION
CR619309
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 888)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 888)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
    Location/Qualifiers
    1..888
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0DC029YK21"
        /tissue_type="Neuroblastoma Cot 25-normalized"
        /plasmid="pCMVSPORT_6"
ORIGIN
Query Match 100.0%; Score 492; DB 3; Length 888;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
    |||
Db 87 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 146

QY 61 CTACCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
    |||
Db 147 CTACCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 206

```

```
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180
Db 207 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 266
QY 181 GGCAGGGCTCAGGAAGGCTGTACAGTGGCTGAGATTCTCAAGCGGGGGTCCAGGC 240
Db 267 GGCAGGGCTCAGGAAGGCTGTACAGTGGCTGAGATTCTCAAGCGGGGGTCCAGGC 326
QY 241 CTCACACAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 300
Db 327 CTCACACAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 386
QY 301 CCTGACACAGGCTAGACCCCTCAGTCGCGGGCCATGTGCTGAGTGGGTGCTG 360
Db 387 CCTGACACAGGCTAGACCCCTCAGTCGCGGGCCATGTGCTGAGTGGGTGCTG 446
QY 361 CTCAGCGGGACCCCTGAGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCT 420
Db 447 CTCAGCGGGACCCCTGAGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCT 506
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTTCCGGAAGAGGGCTCGAGAC 480
Db 507 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTTCCGGAAGAGGGCTCGAGAC 566
QY 481 ACCGATCGTGA 492
Db 567 ACCGATCGTGA 578
```

```
RESULT 7
BQ218097
LOCUS
DEFINITION
5', mRNA sequence.
AGENCOURT 7510498 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6018387
```

```
ACCESSION
BQ218097
VERSION
BQ218097.1 GI:20399497
KEYWORDS
EST.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens (human)
```

```
REFERENCE
BQ218097
LOCUS
DEFINITION
5', mRNA sequence.
AGENCOURT 7510498 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6018387
5', mRNA sequence.
AGENCOURT 8485520 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6305396
```

```
ACCESSION
BQ643214
VERSION
BQ643214.1 GI:21767386
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 918)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13219 row: d column: 04
High quality sequence stop: 539.
```

```
FEATURES
source
1..918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6018387"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_70"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo df.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
```

## ORIGIN

```
Query Match 100.0%; Score 492; DB 5; Length 918;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
```

```
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 60
Db 58 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 117
QY 61 CTACCTCTTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 118 CTACCTCTTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 177
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180
Db 178 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 237
QY 181 GGCAGGGCTGCGAGGAAAGGCTGTACAGTCTCGGCTGAGATTGTCAAGCGGGGGTCCAGGC 240
Db 238 GGCAGGGCTGCGAGGAAAGGCTGTACAGTCTCGGCTGAGATTGTCAAGCGGGGGTCCAGGC 297
QY 241 CTGACACAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 300
Db 298 CTGACACAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 357
QY 301 CCTGACACAGGCTAGACCCCTCAGTCGCGGGCCATGTGCTGAGTGGGTGCTG 360
Db 358 CCTGACACAGGCTAGACCCCTCAGTCGCGGGCCATGTGCTGAGTGGGTGCTG 417
QY 361 CTCAGCGGGACCCCTGAGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 418 CTCAGCGGGACCCCTGAGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 477
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTTCCGGAAGAGGGCTCGAGAC 480
Db 478 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTTCCGGAAGAGGGCTCGAGAC 537
QY 481 ACCGATCGTGA 492
Db 538 ACCGATCGTGA 549
```

```
RESULT 8
BQ643214
LOCUS
DEFINITION
5', mRNA sequence.
AGENCOURT 8485520 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6305396
5', mRNA sequence.
AGENCOURT 8485520 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6305396
5', mRNA sequence.
AGENCOURT 8485520 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6305396
5', mRNA sequence.
AGENCOURT 8485520 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6305396
```

```
ACCESSION
BQ643214
VERSION
BQ643214.1 GI:21767386
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 940)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Ruben Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
```

```
Plate: LLCM2527 row: j column: 21
High quality sequence stop: 615.
FEATURES
source
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6305396"
/tissue_type="lymphoma, cell line"
```

/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGACGAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

## ORIGIN

Query Match 100.0%; Score 492; DB 5; Length 940;  
Best Local Similarity 100.0%; Pred. No. 2.8e-115;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGCACTACCGAAAGCTGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
DB 54 ATGAGCACTACCGAAAGCTGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 113  
QY 61 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 120  
DB 114 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 173  
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
DB 174 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 233  
QY 181 GGCAGGGCTCAGAAAGGCTGTGAGTGGCTGAGATTGTCAAGGGGGGTCGCCAGGC 240  
DB 234 GGCAGGGCTCAGAAAGGCTGTGAGTGGCTGAGATTGTCAAGGGGGGTCGCCAGGC 293  
QY 241 CTCACCAAGCTCAGCAAGCTACGTTTCTTCCATGAGAGCAAGTGGTCCAGCCTCA 300  
DB 294 CTCACCAAGCTCAGCAAGCTACGTTTCTTCCATGAGAGCAAGTGGTCCAGCCTCA 353  
QY 301 CTTGACACAGGGCTAGACCCCTTACAGTGGCGGCGGCTGCTGCTGAGTGGGGTCTG 360  
DB 354 CTTGACACAGGGCTAGACCCCTTACAGTGGCGGCGGCTGCTGCTGAGTGGGGTCTG 413  
QY 361 CTGAGCGGAGCCCTTGAGACCCCAATGAGTGGTTACCAACCCCGAGAGCACCCCT 420  
DB 414 CTGAGCGGAGCCCTTGAGACCCCAATGAGTGGTTACCAACCCCGAGAGCACCCCT 473  
QY 421 GGCCTGGTTTCCATGCCAGCTCCAGCTGGGCTTGGGCTTCCGAGAGAGGGTCCAGAC 480  
DB 474 GGCCTGGTTTCCATGCCAGCTCCAGCTGGGCTTGGGCTTCCGAGAGAGGGTCCAGAC 533  
QY 481 ACCCGATCGTGA 492  
DB 534 ACCCGATCGTGA 545

## RESULT 9

BQ056655  
LOCUS  
DEFINITION AGENCOURT\_6940426 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5809057  
5', mRNA sequence.

## ACCESSION

BQ056655

## VERSION

BQ056655.1 GI:19815982

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1000)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LCM2054 row: b column: 02  
High quality sequence stop: 652.  
Location/Qualifiers  
1. .1000  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5809057"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGACGAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

## FEATURES

source

## ORIGIN

Query Match 100.0%; Score 492; DB 5; Length 1000;  
Best Local Similarity 100.0%; Pred. No. 2.8e-115;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGCACTACCGAAAGCTGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
DB 50 ATGAGCACTACCGAAAGCTGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 109  
QY 61 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 120  
DB 110 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 169  
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
DB 170 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 229  
QY 181 GGCAGGGCTCAGAAAGGCTGTGAGTGGCTGAGATTGTCAAGGGGGTCCAGGC 240  
DB 230 GGCAGGGCTCAGAAAGGCTGTGAGTGGCTGAGATTGTCAAGGGGGTCCAGGC 289  
QY 241 CTGACCAAGCTCACCAGCTAGTTTCTTCCATGAGGAGAGCTGGTCCAGCCTCA 300  
DB 290 CTGACCAAGCTCACCAGCTAGTTTCTTCCATGAGGAGAGCTGGTCCAGCCTCA 349  
QY 301 CTTGACACAGGGCTAGACCCCTTACAGTGGCGGCGGCTGCTGCTGAGTGGGGTCTG 360  
DB 350 CTTGACACAGGGCTAGACCCCTTACAGTGGCGGCGGCTGCTGCTGAGTGGGGTCTG 409  
QY 361 CTCAGCGGAGCCCTTGAGACCCCAATGAGTGGTTACCAACCCCGAGAGCACCCCT 420  
DB 410 CTCAGCGGAGCCCTTGAGACCCCAATGAGTGGTTACCAACCCCGAGAGCACCCCT 469  
QY 421 GGCCTGGTTTCCATGCCAGCTCCAGCTGGGCTTGGGCTTCCGAGAGAGGGTCCAGAC 480  
DB 470 GGCCTGGTTTCCATGCCAGCTCCAGCTGGGCTTGGGCTTCCGAGAGAGGGTCCAGAC 529  
QY 481 ACCCGATCGTGA 492  
DB 530 ACCCGATCGTGA 541

## RESULT 10

BQ056288

LOCUS

DEFINITION AGENCOURT\_6773378 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5808601

BQ056288 1025 bp mRNA linear EST 29-MAR-2002  
AGENCOURT\_6773378 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5808601



```
5', mRNA sequence.
BQ056288
VERSION BQ056288.1 GI:19815628
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2052 row: o column: 02
High quality sequence stop: 555.
Location/Qualifiers
1..1025
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5808601"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
source
FEATURES
SOURCE
1..1074
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3836670"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
Query Match 100.0%; Score 492; DB 5; Length 1025;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGGACACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
DB 54 ATGAGGACACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 113
QY 61 CTACCTCTGATACCTTGAAGTCGGGTCGAGATGGCGAATAATCCCAACCTGCTG 120
DB 114 CTACCTCTGATACCTTGAAGTCGGGTCGAGATGGCGAATAATCCCAACCTGCTG 173
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT 180
DB 174 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT 233
QY 181 GGCAGGGCTCGAGGAAGGCTGTGAGTGTGGCTGTGAGATTGTCAAGCGGGGTCGCCAGGC 240
DB 234 GGCAGGGCTCGAGGAAGGCTGTGAGTGTGGCTGTGAGATTGTCAAGCGGGGTCGCCAGGC 293
QY 241 CTGACACAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 300
DB 294 CTGACACAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 353
QY 301 CCTGACACAGGCTAGACCCCTCCACAGTGGCGGCGGATGTGCTGCGAGTGGGTGCTG 360
DB 354 CCTGACACAGGCTAGACCCCTCCACAGTGGCGGCGGATGTGCTGCGAGTGGGTGCTG 413
QY 361 CTGAGCGGGACCCCTTGAGCCCAATGAGTGTGGTTACCAACCCCGAGGACCCCTT 420
414 CTCAGCGGGACCCCTTGAGCCCAATGAGTGTGGTTACCAACCCCGAGGACCCCTT 473
421 GGCCTGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCGGAAGAGGCTCGAGAC 480
DB 474 GGCCTGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCGGAAGAGGCTCGAGAC 533
481 ACCCGATCGTGA 492
534 ACCCGATCGTGA 545
RESULT 11
BE742621 1074 bp mRNA linear EST 15-SEP-2000
LOCUS 601575430F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3836670 5',
DEFINITION mRNA sequence.
ACCESSION BE742621
VERSION BE742621.1 GI:10156600
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM520 row: k column: 07
High quality sequence stop: 717.
Location/Qualifiers
1..1074
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3836670"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
Query Match 100.0%; Score 492; DB 2; Length 1074;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGGACACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
DB 33 ATGAGGACACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 92
QY 61 CTACCTCTGATACCTTGAAGTCGGGTCGAGATGGCGAATAATCCCAACCTGCTG 120
DB 93 CTACCTCTGATACCTTGAAGTCGGGTCGAGATGGCGAATAATCCCAACCTGCTG 152
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT 180
DB 153 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT 212
QY 181 GGCAGGGCTCGAGGAAGGCTGTGAGTGTGGTGTGGTTACCAACCCCGAGGCGGGTCC 240
```

```

Db      213  GGAGAGGCTGACGAAAGGCTGTGAGTGGCTGAGATTGTCAAGGGGGGTCCTCCAGGC 272
QY      241  CTGACACAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCGAGCCTCA 300
Db      273  CTGACACAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCGAGCCTCA 332
QY      301  CCTGACACAGGCTAGACCCCTCTCACAGTGCAGCCGCGCATGTGCTGCTGAGTGGGTGCTG 360
Db      333  CCTGACACAGGCTAGACCCCTCTCACAGTGCAGCCGCGCATGTGCTGCTGAGTGGGTGCTG 392
QY      361  CTGAGCGGAGACCCCTCTGAGACCCCAATGAGTGGTTACAAACCCGAGAGCACCCCTCT 420
Db      393  CTGAGCGGAGACCCCTCTGAGACCCCAATGAGTGGTTACAAACCCGAGAGCACCCCTCT 452
QY      421  GGCTGGGTTCATGCCAGCTCCAGCTGAGCTGGGCGCTCTGCTCCGAGAGAGGCTCCAGAC 480
Db      453  GGCTGGGTTCATGCCAGCTCCAGCTGAGCTGGGCGCTCTGCTCCGAGAGAGGCTCCAGAC 512
QY      481  ACCCGATCGTGA 492
Db      513  ACCCGATCGTGA 524

```

```

RESULT 12
BQ063771
LOCUS      BQ063771 1079 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6831996 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925418
5', mRNA sequence.
ACCESSION BQ063771
VERSION BQ063771.1 GI:19891815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2097 row: b column: 11
High quality sequence stop: 520.
Location/Qualifiers
1. .1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5925418"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

```

FEATURES
source
Query Match 100.0%; Score 492; DB 5; Length 1079;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
ORIGIN

```

```

Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1  ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGAGCCTTCCCAATGCCCCAG 60
Db      54  ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGAGCCTTCCCAATGCCCCAG 113
QY      61  CTACCTCTCTGATACCTTGTAGATGGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db      114  CTACCTCTCTGATACCTTGTAGATGGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 173
QY      121  GGGTTGGCTCTGGGTCTGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db      174  GGGTTGGCTCTGGGTCTGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 233
QY      181  GGCAGGGCTGCAGGAAGGCTGCTAGCTGCGCTGAGATGTCAAGGGGGGGTCCAGAGC 240
Db      234  GGCAGGGCTGCAGGAAGGCTGCTAGCTGCGCTGAGATGTCAAGGGGGGGTCCAGAGC 293
QY      241  CTGCAACCAAGCTACCAAGCTAGCTTTCTTTCAGACTGAGGACAGCTGGGTCCGAGCCTCA 300
Db      294  CTGCAACCAAGCTACCAAGCTAGCTTTCTTTCAGACTGAGGACAGCTGGGTCCGAGCCTCA 353
QY      301  CTGACACAGGGCTAGACCCCTCTCACAGTGCAGCCGCGCATGTGCTGAGTGGGTGCTG 360
Db      354  CTGACACAGGGCTAGACCCCTCTCACAGTGCAGCCGCGCATGTGCTGAGTGGGTGCTG 413
QY      361  CTCAGCGGGACCCCTCTGAGACCCCAATGAGTGGTTACCAACCCGAGAGCACCCCTCT 420
Db      414  CTCAGCGGGACCCCTCTGAGACCCCAATGAGTGGTTACCAACCCGAGAGCACCCCTCT 473
QY      421  GGCCTGGGTTCATGCCAGCTCCAGCTGCTGCGCTCTGCTCCGAGAGGGCTCCAGAC 480
Db      474  GGCCTGGGTTCATGCCAGCTCCAGCTGCTGCGCTCTGCTCCGAGAGGGCTCCAGAC 533
QY      481  ACCCGATCGTGA 492
Db      534  ACCCGATCGTGA 545

```

RESULT 13  
 BQ06218  
 LOCUS BQ06218 1104 bp mRNA linear EST 05-MAR-2002  
 DEFINITION AGENCOURT\_6554071 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5555733  
 5', mRNA sequence.  
 ACCESSION BQ06218  
 VERSION BQ06218.1 GI:19123029  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgs.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM12276 row: f column: 22  
 High quality sequence stop: 630.  
 Location/Qualifiers  
 1. .1104  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5555733"  
 /tissue\_type="leiomyosarcoma"

/lab host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally... Primer: Oligo dt.  
Average insert size 2.1 kb."

## ORIGIN

Query Match 100.0%; Score 492; DB 4; Length 1104;  
Best Local Similarity 100.0%; Pred. No. 2.9e-115;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGGACATACCGAAAGCTGGCTGTAGAGTCCAGCGCTTCCCAATGCCCCAG 60  
Db 163 ATGAGGACATACCGAAAGCTGGCTGTAGAGTCCAGCGCTTCCCAATGCCCCAG 222  
Qy 61 CTACCTCTGATACCTTGGAGTCCGGGTCCGAGATGGCAGCAAAATTCGCAACTCTG 120  
Db 223 CTACCTCTGATACCTTGGAGTCCGGGTCCGAGATGGCAGCAAAATTCGCAACTCTG 282  
Qy 121 GGGTTGGCTTGGGTTCGGTTGGAGGGCGGAGTCTCGGCATGATGTTCTCAGGTTCT 180  
Db 283 GGGTTGGCTTGGGTTCGGTTGGAGGGCGGAGTCTCGGCATGATGTTCTCAGGTTCT 342  
Qy 181 GGCAGGCTCAGGAAAGCTGTGCTGCTGAGTTCGAGATTCAGCGCGGGTCCAGGC 240  
Db 343 GGCAGGCTCAGGAAAGCTGTGCTGCTGAGTTCGAGATTCAGCGCGGGTCCAGGC 402  
Qy 241 CTGACACAGCTCAGCAAGCTACGTTTCTTCCAGACTGAGGACAGCTGGGTCCAGGCTCA 300  
Db 403 CTGACACAGCTCAGCAAGCTACGTTTCTTCCAGACTGAGGACAGCTGGGTCCAGGCTCA 462  
Qy 301 CTGACACAGGCTAGACCCCTCAGAGTGGCGCCCATGTCGCTGAGTGTGGGTGCTG 360  
Db 463 CTGACACAGGCTAGACCCCTCAGAGTGGCGCCCATGTCGCTGAGTGTGGGTGCTG 522  
Qy 361 CTGACCGGAGCCCTGAGCCCAATGAGTGTGGTTCACCAACCCCGAGGACCCCT 420  
Db 523 CTGACCGGAGCCCTGAGCCCAATGAGTGTGGTTCACCAACCCCGAGGACCCCT 582  
Qy 421 GGCCTGGTTCATGCCAGCTCAGCTGAGTGTGGTTCACCAACCCCGAGGCTCGAGAC 480  
Db 583 GGCCTGGTTCATGCCAGCTCAGCTGAGTGTGGTTCACCAACCCCGAGGCTCGAGAC 642  
Qy 481 ACCCGATCGTGA 492  
Db 643 ACCCGATCGTGA 654

## RESULT 14

BQ070451  
LOCUS  
DEFINITION AGENCOURT 6839400 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:5922837  
5', mRNA sequence.

## ACCESSION

BQ070451

## VERSION

BQ070451.1 GI:19899497

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LLCW2090 row: f column: 22

High quality sequence stop: 678.

## FEATURES

## source

1..967  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5922837"  
/issue\_type="neuroblastoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_47"  
/note="Organ: Brain; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 99.8%; Score 491; DB 5; Length 967;  
Best Local Similarity 99.8%; Pred. No. 5.1e-115;  
Matches 491; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGGACATACCGAAAGCTGGCTGTAGAGTCCAGCGCTTCCCAATGCCCCAG 60  
Db 254 ATGAGGACATACCGAAAGCTGGCTGTAGAGTCCAGCGCTTCCCAATGCCCCAG 313  
Qy 61 CTACCTCTGATACCTTGGAGTCCGGGTCCGAGATGGCAGCAAAATTCGCAACTCTG 120  
Db 314 CTACCTCTGATACCTTGGAGTCCGGGTCCGAGATGGCAGCAAAATTCGCAACTCTG 373  
Qy 121 GGGTTGGCTTGGGTTCGGTTGGAGGGCGGAGTCTCGGCATGATGTTCTCAGGTTCT 180  
Db 374 GGGTTGGCTTGGGTTCGGTTGGAGGGCGGAGTCTCGGCATGATGTTCTCAGGTTCT 433  
Qy 181 GGCAGGGCTGACGAAAGCTGTGCTGAGTTCAGAGTTCAGCGGGTCCAGGC 240  
Db 434 GGCAGGGCTGACGAAAGCTGTGCTGAGTTCAGAGTTCAGCGGGTCCAGGC 493  
Qy 241 CTGACACAGCTCAGCAAGCTACGTTTCTTCCAGACTGAGGACAGCTGGGTCCAGGCTCA 300  
Db 494 CTGACACAGCTCAGCAAGCTACGTTTCTTCCAGACTGAGGACAGCTGGGTCCAGGCTCA 553  
Qy 301 CTGACACAGGGCTAGACCCCTCAGAGTCCGCGCCCATGTCGCTGAGTGTGGGTGCTG 360  
Db 554 CTGACACAGGGCTAGACCCCTCAGAGTCCGCGCCCATGTCGCTGAGTGTGGGTGCTG 613  
Qy 361 CTGACCGGAGCCCTGAGCCCAATGAGTGTGGTTCACCAACCCCGAGGACCCCT 420  
Db 614 CTGACCGGAGCCCTGAGCCCAATGAGTGTGGTTCACCAACCCCGAGGACCCCT 673  
Qy 421 GGCCTGGTTCATGCCAGCTCAGCTGAGTGTGGTTCGCGAAGAGGGCTCGAGAC 480  
Db 674 GGCCTGGTTCATGCCAGCTCAGCTGAGTGTGGTTCGCGAAGAGGGCTCGAGAC 733  
Qy 481 ACCCGATCGTGA 492  
Db 734 ACCCGATCGTGA 745

## RESULT 15

## CB110489

## LOCUS

## DEFINITION

K-EST0151955 L5HLK1 Homo sapiens cDNA clone L5HLK1-5-D07 5', mRNA

## ACCESSION

CB110489

## VERSION

CB110489.1 GI:27936296

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 604)  
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
 21C Frontier Korean EST Project 2001  
 Unpublished (2002)  
 Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongu@kribb.re.kr  
 Plate: 5 row: D column: 07  
 High quality sequence stop: 604.

Location/Qualifiers  
 1. 604

Source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="J5HLK1-5-D07"  
 /sex="M"  
 /cell\_line="HLK-1"  
 /lab\_host="Top10F"  
 /clone\_lib="J5HLK1"  
 /notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

#### FEATURES

Source

1. 604

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="J5HLK1-5-D07"

/sex="M"

/cell\_line="HLK-1"

/lab\_host="Top10F"

/clone\_lib="J5HLK1"

/notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

/lab\_host="Top10F"

/clone\_lib="J5HLK1"

/notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

/lab\_host="Top10F"

/clone\_lib="J5HLK1"

/notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

/lab\_host="Top10F"

/clone\_lib="J5HLK1"

/notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

/lab\_host="Top10F"

/clone\_lib="J5HLK1"

/notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

/lab\_host="Top10F"

/clone\_lib="J5HLK1"

/notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

#### ORIGIN

Query Match 99.7%; Score 490.4; DB 6; Length 604;  
 Best Local Similarity 99.8%; Pred. No. 6.9e-115;  
 Matches 491; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
 Db 69 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 128  
 QY 61 CTACCTCTCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGTG 120  
 Db 129 CTACCTCTCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGTG 188  
 QY 121 GGGTTCGGCTCTGGGTGGGTGGAGGGGGGCGAGTGTCCGCATGTAGTGTCTCAGGTTCT 180  
 Db 189 GGGTTCGGCTCTGGGTGGGTGGAGGGGGGCGAGTGTCCGCATGTAGTGTCTCAGGTTCT 248  
 QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGGGTGCCAGGC 240  
 Db 249 GGCAGGGCTGCAGGAAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGGGTGCCAGGC 308  
 QY 241 CTGCACAGCTACCAAGCTAGCTTTCTTCAGCTGAGGACAGCTGGGTCCAGCTCA 300  
 Db 309 CTGCACAGCTACCAAGCTAGCTTTCTTCAGCTGAGGACAGCTGGGTCCAGCTCA 368  
 QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGGCGCCCATGTGCTGCACTGGGTGGGTG 360  
 Db 369 CCAGACACAGGGCTAGACCCCTCACAGTGGCGCCCATGTGCTGCACTGGGTGGGTG 428

QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGTTACCAACCCCGAGGAGACCCCT 420  
 Db 429 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGTTACCAACCCCGAGGAGACCCCT 488  
 QY 421 GGCCTGGGTTCATGCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480  
 Db 489 GGCCTGGGTTCATGCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 548  
 QY 481 ACCCGATCGTGA 492  
 Db 549 ACCCGATCGTGA 560

Search completed: January 31, 2005, 18:07:49  
 Job time : 2438 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 17:27:15 ; Search time 423 Seconds

(without alignments)  
6683.161 Million cell updates/sec

Title: US-10-057-813-13

Perfect score: 492

Sequence: 1 atggagcactaccggaagc.....ctcgagacacccgatcgta 492

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUE\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	492	100.0	894	9	US-09-731-872-218
2	492	100.0	894	10	US-09-876-997-218
3	484	98.4	884	11	US-09-978-360A-311
4	373.8	76.0	469	10	US-09-918-999-21099
c 5	368.4	74.9	1168	9	US-09-981-353-1
c 6	368.4	74.9	1168	15	US-10-084-817-60
c 7	223.8	45.5	560	16	US-10-152-319A-920
8	197.8	40.2	299	18	US-10-425-115-52678
9	74.6	15.2	1941	15	US-10-037-270-5
10	74.6	15.2	1941	15	US-10-117-722-5
11	68.2	13.9	412	9	US-09-983-965-2039
12	40.4	8.2	929	18	US-10-653-047-7620
					Sequence 218, App
					Sequence 218, App
					Sequence 311, App
					Sequence 21099, A
					Sequence 1, Appli
					Sequence 60, Appl
					Sequence 920, App
					Sequence 52678, A
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 2039, Ap
					Sequence 7620, Ap

c 13	39.4	8.0	32189	9	US-09-764-877-3291	Sequence 3291, Ap
c 14	39.4	8.0	32189	16	US-10-242-515-3291	Sequence 3291, Ap
c 15	38.8	7.9	972	15	US-10-156-761-4789	Sequence 4789, Ap
c 16	38.8	7.9	9025608	15	US-10-156-761-1	Sequence 1, Appli
c 17	37.8	7.7	594	14	US-10-123-155-10	Sequence 10, Appl
c 18	37.8	7.7	594	15	US-10-146-731-10	Sequence 10, Appl
c 19	37.8	7.7	594	15	US-10-140-472-10	Sequence 10, Appl
c 20	37.8	7.7	594	15	US-10-141-761-10	Sequence 10, Appl
c 21	37.8	7.7	594	15	US-10-142-885-10	Sequence 10, Appl
c 22	37.8	7.7	594	15	US-10-158-790-10	Sequence 10, Appl
c 23	37.8	7.7	594	15	US-10-137-871-10	Sequence 10, Appl
c 24	37.8	7.7	594	15	US-10-140-923-10	Sequence 10, Appl
c 25	37.8	7.7	594	15	US-10-141-756-10	Sequence 10, Appl
c 26	37.8	7.7	594	15	US-10-141-759-10	Sequence 10, Appl
c 27	37.8	7.7	594	15	US-10-140-805-10	Sequence 10, Appl
c 28	37.8	7.7	594	15	US-10-140-864-10	Sequence 10, Appl
c 29	37.8	7.7	594	16	US-10-142-426-10	Sequence 10, Appl
c 30	37.4	7.6	645	15	US-10-156-761-5727	Sequence 5727, Ap
c 31	37.4	7.6	9025608	15	US-10-156-761-1	Sequence 1, Appli
c 32	37.2	7.6	1218	16	US-10-425-114-21994	Sequence 21994, A
c 33	37.2	7.6	1520	18	US-10-425-115-16813	Sequence 16813, A
c 34	37.2	7.6	46819	14	US-10-114-170-72	Sequence 72, Appl
c 35	36.6	7.4	1593	15	US-10-156-761-1221	Sequence 1221, Ap
c 36	36.2	7.4	500	10	US-09-770-961-441	Sequence 441, App
c 37	36	7.3	1363	18	US-10-842-005A-2	Sequence 2, Appli
c 38	36	7.3	11100	15	US-10-329-079-5	Sequence 5, Appli
c 39	36	7.3	13315	15	US-10-329-079-1	Sequence 1, Appli
c 40	35.6	7.2	831	15	US-10-156-761-4034	Sequence 4034, Ap
c 41	35.6	7.2	22306	14	US-10-114-170-251	Sequence 251, App
c 42	35.4	7.2	493	9	US-09-783-590-3619	Sequence 3619, Ap
c 43	35.4	7.2	1993	18	US-10-425-115-79876	Sequence 79876, A
c 44	35.4	7.2	2662	9	US-09-925-301-2	Sequence 2, Appli
c 45	35.4	7.2	4257	9	US-09-825-288A-1	Sequence 1, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-731-872-218  
; Sequence 218, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 218  
; LENGTH: 894  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 78..566  
US-09-731-872-218

Query Match 100.0%; Score 492; DB 9; Length 894;

Best Local Similarity 100.0%; Pred. No. 3.4e-140;

Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCACTACCGGAAGCTGGCTCTGTAGACTCCAGCGCTTCCCAATGCCCCAG 60

Db 78 ATGAGCACTACCGGAAGCTGGCTCTGTAGACTCCAGCGCTTCCCAATGCCCCAG 137

QY 61 CTACCTCTGATACCTTGGATGCGGGTCCGAGTGGCAGCAAAATTCGCAACCTGCTG 120  
|  
Db 138 CTACCTCTGATACCTTGGATGCGGGTCCGAGTGGCAGCAAAATTCGCAACCTGCTG 197  
|  
QY 121 GGGTTGGCTCTGGGTCCGGTTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
|  
Db 198 GGGTTGGCTCTGGGTCCGGTTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 257  
|  
QY 181 GGCAGGGCTCCAGGAAGGCTGTGAGTGGCTGGCTGAGATTGTCAAGCGGCGGGTCCAGGC 240  
|  
Db 258 GGCAGGGCTCCAGGAAGGCTGTGAGTGGCTGGCTGAGATTGTCAAGCGGCGGGTCCAGGC 317  
|  
QY 241 CTGACACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGCAGCTGGGTCCAGGCTCA 300  
|  
Db 318 CTGACACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGCAGCTGGGTCCAGGCTCA 377  
|  
QY 301 CTTGACACAGGGTAGACCCCTTACAGTGGCGGCGCATGTGCTGCGAGTGTGGGTGCTG 360  
|  
Db 378 CTTGACACAGGGTAGACCCCTTACAGTGGCGGCGCATGTGCTGCGAGTGTGGGTGCTG 437  
|  
QY 361 CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCT 420  
|  
Db 438 CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCT 497  
|  
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480  
|  
Db 498 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 557  
|  
QY 481 ACCGATCGTGA 492  
|  
Db 558 ACCGATCGTGA 569  
|

## RESULT 2

US-09-876-997-218  
; Sequence 218, Application US/09876997  
; Publication No. US20030152921A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US4.CIP  
; CURRENT APPLICATION NUMBER: US/09/876,997  
; CURRENT FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 09/731,872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 218  
; LENGTH: 894  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 78..566  
US-09-876-997-218

Query Match 100.0%; Score 492; DB 10; Length 894;  
Best Local Similarity 100.0%; Pred. No. 3.4e-140;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGAAAGCTGCTGTAGAGTCCAGCGCTTCCCAATGCCCGAG 60  
|  
Db 78 ATGAGCACTACCGAAAGCTGCTGTAGAGTCCAGCGCTTCCCAATGCCCGAG 137  
|  
QY 61 CTACCTCTGATACCTTGGATGCGGGTCCGAGTGGCAGCAAAATTCGCAACCTGCTG 120  
|  
Db 138 CTACCTCTGATACCTTGGATGCGGGTCCGAGTGGCAGCAAAATTCGCAACCTGCTG 197  
|

QY 121 GGGTTGGCTCTGGGTCCGGTTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
|  
Db 198 GGGTTGGCTCTGGGTCCGGTTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 257  
|  
QY 181 GGCAGGGCTCCAGGAAGGCTGTGAGTGGCTGGCTGAGATTGTCAAGCGGCGGGTCCAGGC 240  
|  
Db 258 GGCAGGGCTCCAGGAAGGCTGTGAGTGGCTGGCTGAGATTGTCAAGCGGCGGGTCCAGGC 317  
|  
QY 241 CTGACACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGCAGCTGGGTCCAGGCTCA 300  
|  
Db 318 CTGACACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGCAGCTGGGTCCAGGCTCA 377  
|  
QY 301 CTTGACACAGGGTAGACCCCTTACAGTGGCGGCGCATGTGCTGCGAGTGTGGGTGCTG 360  
|  
Db 378 CTTGACACAGGGTAGACCCCTTACAGTGGCGGCGCATGTGCTGCGAGTGTGGGTGCTG 437  
|  
QY 361 CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCT 420  
|  
Db 438 CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCT 497  
|  
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480  
|  
Db 498 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 557  
|  
QY 481 ACCGATCGTGA 492  
|  
Db 558 ACCGATCGTGA 569  
|

## RESULT 3

US-09-978-360A-311  
; Sequence 311, Application US/09978360A  
; Publication No. US20040110939A1  
; GENERAL INFORMATION:  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; APPLICANT: Clusel, Catherine  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
; FILE REFERENCE: 56.US4.CIP  
; CURRENT APPLICATION NUMBER: US/09/978,360A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: US 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: US 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: US 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: US 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: US 60/099,273  
; PRIOR FILING DATE: -09-04  
; PRIOR APPLICATION NUMBER: US 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: US 09/215,435  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: PCT/IB98/02122  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: US 09/247,155  
; PRIOR FILING DATE: 1999-02-09  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 810  
; SOFTWARE: Patent.pm  
; SEQ ID NO 311  
; LENGTH: 884  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS

```
; LOCATION: 69..557
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 69..224
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.6999980926514
; OTHER INFORMATION: seq LGLALGRLEGSARH
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 849..854
; NAME/KEY: polyA_site
; LOCATION: 870..883
; US-09-978-360A-311

Query Match          98.4%; Score 484; DB 11; Length 884;
Best Local Similarity 97.8%; Pred. No. 9.3e-138;
Matches 481; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCCAATGCCCCAG 60
Db 69 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCCAATGCCCCAG 128
QY 61 CTACCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 129 CTACCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 188
QY 121 GGGTTGGCTCTGGGTGGGTTGGAGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180
Db 189 GGGTTGGCTCTGGGTGGGTTGGAGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 248
QY 181 GGCAGGGCTGCGAGAAAGCTGTGAGTGTGGTGTGAGATTTGCAAGCGGCGGGTCCAGGC 240
Db 249 GGCAGGGCTGCGAGAAAGCTGTGAGTGTGGTGTGAGATTTGCAAGCGGCGGGTCCAGGC 308
QY 241 CTCACCACTACCAAGCTACGTTTCTTCAACTGAGACAGCTGGGTCCAGCCTCA 300
Db 309 CTCACCACTACCAAGCTACGTTTCTTCAACTGAGACAGCTGGGTCCAGCCTCA 368
QY 301 CTGACACAGGGCTAGACCCCTTACAGTGTGGCGGCGCATGTGCTGCTGAGTGTGGGTGCTG 360
Db 369 CTGACACAGGGCTARACCCCTTACAGTGTGGCGGCGCATGTGCTGCTGAGTGTGGGTGCTG 428
QY 361 CTCAGCGGACCCCTTGACCCCAATGAGTGTGGTTACCAACCCCGGAGCAGCCCTT 420
Db 429 CTCAGCGGACCCCTTGACCCCAATGAGTGTGGTTACCAACCCCGGAGCAGCCCTT 488
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGAGTGTGGCCCTGTTCCGAAAGAGGGCTCGAGC 480
Db 489 GGCCTGGTTCCATGCCAGCTCCAGCTGAGTGTGGCCCTGTTCCGAAAGAGGGCTCGARAC 548
QY 481 ACCCGATCGTGA 492
Db 549 ACCCGATCGTGA 560

RESULT 4
US-09-918-995-21099
; Sequence 21099, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21099
; LENGTH: 469

; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-918-995-21099

Query Match          76.0%; Score 373.8; DB 10; Length 469;
Best Local Similarity 99.2%; Pred. No. 4e-104;
Matches 386; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCCAATGCCCCAG 60
Db 81 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCCAATGCCCCAG 140
QY 61 CTACCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 141 CTACCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 200
QY 121 GGGTTGGCTCTGGGTGGGTTGGAGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGAGTTC 179
Db 201 GGGTTGGCTCTGGGTGGGTTGGAGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGAGTTC 260
QY 180 TGGCAGGGCTGCGAGAAAGCTGTGTGAGTGTGGTGTGAGATTTGTCAAGCGGCGGGTCCCAGG 239
Db 261 TGGCAGGGCTGCGAGAAAGCTGTGTGAGTGTGGTGTGAGATTTGTCAAGCGGCGGGTCCCAGG 320
QY 240 CTGCAACCACTGACCAAGCTAGTGTCTTTCAGACTGAGGACAGCTGGGTCCAGCCTC 299
Db 321 CTGCAACCACTGACCAAGCTAGTGTCTTTCAGACTGAGGACAGCTGGGTCCAGCCTC 380
QY 300 ACTGTACACAGGGCTAGACCCCTTACAGTGTGGCGGCGCATGTGCTGCAAGTGTGGGTGCT 359
Db 381 ACTGTACACAGGGCTAGACCCCTTACAGTGTGGCGGCGCATGTGCTGCAAGTGTGGGTGCT 440
QY 360 GCTCACCGGGACCCCTTGGACCCCAATG 388
Db 441 GCTCATCCGGGACCCCTTGGACCCCAATG 469

RESULT 5
US-09-981-353-1/c
; Sequence 1, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 1168
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 184081.24
; US-09-981-353-1

Query Match          74.9%; Score 368.4; DB 9; Length 1168;
Best Local Similarity 99.7%; Pred. No. 1.9e-102;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 123 GTTGGCTCTGGGTCTGGAGGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCTGG 182
Db 871 GCTGGCTCTGGGTCTGGAGGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCTGG 812
QY 183 CAGGGCTGCGAGAAAGCTGTGAGTGTGGTGTGAGATTTGTCAAGCGGCGGGTCCAGGCT 242
Db 811 CAGGGCTGCGAGAAAGCTGTGAGTGTGGTGTGAGATTTGTCAAGCGGCGGGTCCAGGCT 752
QY 243 GCACCACTGACCAAGCTAGTGTCTTTCAGACTGAGGACAGCTGGGTCCAGCCTCACC 302
```

Db 751 GCACACAGCTCACCAAGTACGTTTCCTTCAGACTGAGAGCAGCTGGGTCCACGCTCAC 692  
Qy 303 TGACACAGGGCTAGACCCCTCTACAGTGGCGCCCATGTGCTGACAGTGGGTGCTGCT 362  
Db 691 TGACACAGGGCTAGACCCCTCTACAGTGGCGCCCATGTGCTGACAGTGGGTGCTGCT 632  
Qy 363 CAGCCGGGACCCCTCTGAGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCTCTG 422  
Db 631 CAGCCGGGACCCCTCTGAGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCTCTG 572  
Qy 423 CTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCTGTTCCCGAAGAGGGCTCGAGAC 482  
Db 571 CTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCTGTTCCCGAAGAGGGCTCGAGAC 512  
Qy 483 CCGATCGTGA 492  
Db 511 CCGATCGTGA 502

## RESULT 6

US-10-084-817-60/c  
; Sequence 60, Application US/10084817  
; Publication No. US20030119009A1  
; GENERAL INFORMATION:  
; APPLICANT: Susan Stuart  
; APPLICANT: Jed G. Nuchtern  
; APPLICANT: Sharon E. Plon  
; APPLICANT: Jason M. Shonet  
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
; FILE REFERENCE: PA-0046 US  
; CURRENT APPLICATION NUMBER: US/10/084,817  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/270,784  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 365  
; SOFTWARE: PERL Program  
; SEQ ID NO 60  
; LENGTH: 1168  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030119009A1 184081.24  
US-10-084-817-60

Query Match 74.9%; Score 368.4; DB 15; Length 1168;  
Best Local Similarity 99.7%; Pred. No. 1.9e-102;  
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 123 GTTGGCTCTGGGTCTGGTGGAGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCTGG 182  
Db 871 GCTGGCTCTGGGTCTGGTGGAGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCTGG 812  
Qy 183 CAGGGCTGCAGGAAGGCTGTCTAGCTGCGCTGAGATTGTCAAGCGCGGGTCCAGGCT 242  
Db 811 CAGGGCTGCAGGAAGGCTGTCTAGCTGCGCTGAGATTGTCAAGCGCGGGTCCAGGCT 752  
Qy 243 GCACAGCTCAACAAGCTAGCTTTCTTCACTGAGGACAGCTGGGTCCAGGCTCAC 302  
Db 751 GCACAGCTCACCAAGCTAGCTTTCTTCACTGAGGACAGCTGGGTCCAGGCTCAC 692  
Qy 303 TGACACAGGCTAGACCCCTCTACAGTGGCGCCCATGTGCTGCAAGTGGGTGCTGCT 362  
Db 691 TGACACAGGCTAGACCCCTCTACAGTGGCGCCCATGTGCTGCAAGTGGGTGCTGCT 632  
Qy 363 CAGCCGGGACCCCTCTGAGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCTCTG 422  
Db 631 CAGCCGGGACCCCTCTGAGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCTCTG 572  
Qy 423 CTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCTGTTCCCGAAGAGGGCTCGAGAC 482  
Db 571 CTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCTGTTCCCGAAGAGGGCTCGAGAC 512

Qy 483 CCGATCGTGA 492  
Db 511 CCGATCGTGA 502  
RESULT 7  
US-10-152-319A-920/c  
; Sequence 920, Application US/10152319A  
; Publication No. US20040072160A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Higgs, Brandon  
; APPLICANT: Casle, Arthur  
; APPLICANT: Elashoff, Michael  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5089-US  
; CURRENT APPLICATION NUMBER: US/10/152,319A  
; CURRENT FILING DATE: 2002-05-22  
; PRIOR APPLICATION NUMBER: US 60/292,335  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/297,523  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,925  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,810  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/303,807  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/303,808  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/315,047  
; PRIOR FILING DATE: 2001-08-28  
; PRIOR APPLICATION NUMBER: US 60/324,928  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US 60/330,867  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: US 60/330,462  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2221  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 920  
; LENGTH: 560  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. AI104296  
US-10-152-319A-920

Query Match 45.5%; Score 223.8; DB 16; Length 560;  
Best Local Similarity 85.6%; Pred. No. 2.7e-58;  
Matches 249; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
Qy 202 GTCAGTGGCTGAGATTGTCAAGCGCGGGTCCAGGCTGCACAGCTCACCAAGCTA 261  
Db 560 GTCAGTGGCTGAGATTGTCAAGCGCGGGTCCAGGCTGCACAGCTCACCAAGCTT 501  
Qy 262 CGTTCTCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCACCTGACACAGGGCTAGACCC 321  
Db 500 CGTTCTCTTTCAGACTGAGGACAGCTGGGTCCAGGCTCACCTGACACAGGGCTAGACCC 441  
Qy 322 CTCAGAGTGGCGGCGCATGTGCTGAGTGGGTGCTGCTCAGCGGAGACCCCTGGAC 381  
Db 440 CTCAGAGTGGCGGCGCATGTGCTGAGTGGGTGCTGCTCAGCGGAGACCCCTGGAC 381  
Qy 382 CCCAATGAGTGGTTTACCAACCCCGAGGACACCCCTGGCTGGGTTCATGCCAGC 441  
Db 380 CCCAATGAGTGGTTTACCAACCCCGAGGACACCCCTGGCTGGGTTCATGCCAGC 321  
Qy 442 TCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGACACCCGATCGTGA 492



```

; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 5
; LENGTH: 1941
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (138)..(773)
US-10-037-270-5

Query Match          15.2%; Score 74.6; DB 15; Length 1941;
Best Local Similarity 57.5%; Pred. No. 1.1e-12;
Matches 134; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 36 CCCAGCGCCTTCCCAATGCCCCAGCTACCTCTGATACCTTTGAGATGCGGGTCCGAGA 95
Db 239 CCCGGGCTCGGGCCCGTTTCGACACTTGGCGCCGGGGCGGTGCACATCGGGTCAAGGA 298

QY 96 TGCAGACAAATTCGCAACCTGCTGGGGTTGGCTCTGGGTTCGGTTGGAGGGCGGCAGTGC 155
Db 299 AGGCAGCAAGATCGGAACTTGATGCGCTTCGCCACCGCCAGCATGGCGCAGCCAGCCAC 358

QY 156 TCGGCATGTAGTTTCTCAGGTTCTGGCAGGGCTGCAGGAAAGGCTGTACAGTTCGCGTGA 215
Db 359 GCGCGCCATCGTCTTCAGCGGCTTCGCGCGGGCCACCCACCAAAACCTGCACGTGCGCCGA 418

QY 216 GATTGTCAAGCGCGGGTCCCGAGGCTGCACACAGCTCACCAAGCTACGTTTCC 268
Db 419 GATCCTCAAGCGCGGCTTCGCGGGGCTGCACACAGGTCAACCGGCTGCGCTACC 471

RESULT 10
US-10-117-722-5
; Sequence 5, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 5
; LENGTH: 1941
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (138)..(773)
US-10-117-722-5

Query Match          15.2%; Score 74.6; DB 15; Length 1941;
Best Local Similarity 57.5%; Pred. No. 1.1e-12;
Matches 134; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

```

QY 36 CCAGGCGCTTCCCAATGCCAGCTACCTCCTGATACCTTTCAGATCGGGTCCGAGA 95  
|||  
Db 239 CCGGGCTCGGCGCGCTTCGACACCTGGCGCGGCGGTGCACATCGGGTCAAGGA 298  
|||  
QY 96 TGGCAGCAAAATTCGAACCTGCTGGGTTGGCTCTGGGTGCGTTGGAGGCGGCGAGTGC 155  
|||  
Db 299 AGGCAGCAGATCCGGAACCTGATGCTTCGCCACCGCCAGCATGGCGAGCCAC 358  
|||  
QY 156 TGGCATGTAGTGTCTCTCAGGTTCTGCGAGGGCTGCGAGAAAGGCTGAGCTGCGCTGA 215  
|||  
Db 359 GCGGCGCATGCTTTCAGGGCTGCGCGGCGGTGCACCAAAACCGTCAGTGGCGCGGA 418  
|||  
QY 216 GATTGTCAAGCGCGGGTCCAGGCTGCACAGCTGCACCAAGCTACGTTTCC 268  
|||  
Db 419 GATCCTCAAGCGCGCTGCGGGCTGCACAGGTACCGGGCTGGCTACC 471  
|||

## RESULT 11

US-09-983-965-2039  
; Sequence 2039, Application US/09983965  
; Patent No. US20020137160A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
; FILE REFERENCE: 37-21(10297)C  
; CURRENT APPLICATION NUMBER: US/09/983,965  
; PRIOR FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 09/465,231  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/113,678  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 5912  
; SEQ ID NO 2039  
; LENGTH: 412  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (358)  
; OTHER INFORMATION:  
; OTHER INFORMATION: Clone ID: 26-LIB3057-021-Q1-K1-G5  
US-09-983-965-2039

Query Match 13.9%; Score 68.2; DB 9; Length 412;  
Best Local Similarity 78.1%; Pred. No. 8.8e-11;  
Matches 82; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
QY 388 GAGTGTGGTTACCAACCCAGGAGCAGCACCCTGGGCTGGTTCATGCCAGCTCCAGC 447  
|||  
Db 1 GAGTATGGCTTACCAACCCCTGGGCAACCCCTGACCTCGGGCCACACCTGCTGCCAGC 60  
|||  
QY 448 TGTGGCCCTCGTTCCCGAAGAGGGTTCGAGACACCCGATCGTGA 492  
|||  
Db 61 TGTGGCCCTCAGCTAGAGAGAGGGTTCGAGATACCCGGTTCTGA 105  
|||

## RESULT 12

US-10-653-047-7620  
; Sequence 7620, Application US/10653047  
; Publication No. US20040229367A1  
; GENERAL INFORMATION:  
; APPLICANT: Randy M. Berka  
; APPLICANT: Michael W. Rey  
; APPLICANT: Jeffrey R. Shuster  
; APPLICANT: Sakari Kauppinen  
; APPLICANT: Ib Groth Clausen  
; APPLICANT: Peter Bjørke Olsen  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; TITLE OF INVENTION: Expression

FILE REFERENCE: 5849,200-US  
; CURRENT APPLICATION NUMBER: US/10/653,047  
; CURRENT FILING DATE: 2003-08-29  
; PRIOR APPLICATION NUMBER: US/09/533,559  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 09/273,623  
; PRIOR FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 7860  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7620  
; LENGTH: 929  
; TYPE: DNA  
; ORGANISM: Tricodexma reesei  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)....(929)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-653-047-7620

Query Match 8.2%; Score 40.4; DB 18; Length 929;  
Best Local Similarity 54.8%; Pred. No. 0.029;  
Matches 80; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
QY 279 GGACAGCTGGGTCCAGCCTCACCTGACAGGGCTAGACCCCTTCACAGTGGCGGCCA 338  
|||  
Db 305 GCATTGCTGTGCTCGCCCTCGCCCAACCAAGTCCAGGTCAACGCCACACGCAAGTGGTCCGCA 364  
|||  
QY 339 TGTGCTGCAAGTGTGGGTGCTGCTCAGCGGGAGCCCTTGGACCCCATGAGTGTGGTTA 398  
|||  
Db 365 AGTGGCCCCCAGGGCAACGGCAGCGTGGCCAGACGAGATCTACAAGACGTGCATTGACA 424  
|||  
QY 399 CCAAGCCCCCAGGAGCAGCACCCTTGGCC 424  
|||  
Db 425 AGTGATCAACGACCACTACTTCGTC 450  
|||

## RESULT 13

US-09-764-877-3291/c  
; Sequence 3291, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3291  
; LENGTH: 32189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-3291

Query Match 8.0%; Score 39.4; DB 9; Length 32189;  
Best Local Similarity 47.1%; Pred. No. 0.078;  
Matches 121; Conservative 0; Mismatches 136; Indels 0; Gaps 0;  
QY 191 CAGGAAAGGCTGTACAGTCGCTGAGATTGTCAAGCGGCGGTCCAGGCTGCACACAGC 250  
|||  
Db 25559 CAAGGAAGCCAGGGTCCCGGACCACTTCTGCAACCCAGGGGTGTGAGGTGCACCTGC 25500  
|||  
QY 251 TCACCAAGCTAGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGCCTCACCTGCACACAG 310  
|||  
Db 25499 TGCCCAAGGAAACCTGAGTTGAGAAGGTGGCGCCCTGCCCATGGAAGACAGGCCAG 25440  
|||  
QY 311 GGCTAGACCCCTCACAGTGCGCCCATGTCCCTGCAAGTGTGGGTGTGCTCAGCCGGG 370  
|||  
Db 25439 AGGTGGCTGGGCACACACTGCAGAGAGTTGCAAGAACAAAGTTTCTGGTGGCTTGCACAG 25380  
|||  
QY 371 ACCCCCTGAGACCCCAATGAGTGTGTTACCAACCCCGAGGACACCCCTGCGCTGGGTT 430  
|||



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 12:32:30 ; Search time 2949 Seconds  
(without alignments)  
6765.348 Million cell updates/sec

Title: US-10-057-813-13

Perfect score: 492

Sequence: 1 atggagcattaccggaagc.....ctcgagacaccgcgtcgtga 492

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 44308572 seqs, 20275418765 residues

Total number of hits satisfying chosen parameters: 88617144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:\*

- 1: /cgn2\_6/ptodata/1/pna/PCTUS1 COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pna/PCTUS2 COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pna/US06 COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pna/US07 COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pna/US08 COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pna/US081 COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pna/US082 COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pna/US083 COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pna/US084 COMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pna/US085 COMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pna/US086 COMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pna/US087 COMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pna/US088 COMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pna/US089 COMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pna/US090 COMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pna/US091 COMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pna/US092A COMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pna/US092B COMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pna/US093A COMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pna/US093B COMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pna/US094 COMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pna/US095A COMB.seq:\*
- 23: /cgn2\_6/ptodata/1/pna/US095B COMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pna/US095C COMB.seq:\*
- 25: /cgn2\_6/ptodata/1/pna/US095D COMB.seq:\*
- 26: /cgn2\_6/ptodata/1/pna/US096A COMB.seq:\*
- 27: /cgn2\_6/ptodata/1/pna/US096B COMB.seq:\*
- 28: /cgn2\_6/ptodata/1/pna/US096C COMB.seq:\*
- 29: /cgn2\_6/ptodata/1/pna/US096D COMB.seq:\*
- 30: /cgn2\_6/ptodata/1/pna/US096E COMB.seq:\*
- 31: /cgn2\_6/ptodata/1/pna/US097A COMB.seq:\*
- 32: /cgn2\_6/ptodata/1/pna/US097B COMB.seq:\*
- 33: /cgn2\_6/ptodata/1/pna/US097C COMB.seq:\*
- 34: /cgn2\_6/ptodata/1/pna/US098A COMB.seq:\*
- 35: /cgn2\_6/ptodata/1/pna/US098B COMB.seq:\*
- 36: /cgn2\_6/ptodata/1/pna/US098C COMB.seq:\*
- 37: /cgn2\_6/ptodata/1/pna/US098D COMB.seq:\*
- 38: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*
- 39: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*
- 40: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*
- 41: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*
- 42: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*
- 43: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

- 44: /cgn2\_6/ptodata/1/pna/US099G COMB.seq:\*
- 45: /cgn2\_6/ptodata/1/pna/US100A COMB.seq:\*
- 46: /cgn2\_6/ptodata/1/pna/US100B COMB.seq:\*
- 47: /cgn2\_6/ptodata/1/pna/US101A COMB.seq:\*
- 48: /cgn2\_6/ptodata/1/pna/US101B COMB.seq:\*
- 49: /cgn2\_6/ptodata/1/pna/US102A COMB.seq:\*
- 50: /cgn2\_6/ptodata/1/pna/US102B COMB.seq:\*
- 51: /cgn2\_6/ptodata/1/pna/US103A COMB.seq:\*
- 52: /cgn2\_6/ptodata/1/pna/US103B COMB.seq:\*
- 53: /cgn2\_6/ptodata/1/pna/US104A COMB.seq:\*
- 54: /cgn2\_6/ptodata/1/pna/US104B COMB.seq:\*
- 55: /cgn2\_6/ptodata/1/pna/US105A COMB.seq:\*
- 56: /cgn2\_6/ptodata/1/pna/US105B COMB.seq:\*
- 57: /cgn2\_6/ptodata/1/pna/US106A COMB.seq:\*
- 58: /cgn2\_6/ptodata/1/pna/US107A COMB.seq:\*
- 59: /cgn2\_6/ptodata/1/pna/US107B COMB.seq:\*
- 60: /cgn2\_6/ptodata/1/pna/US107C COMB.seq:\*
- 61: /cgn2\_6/ptodata/1/pna/US107D COMB.seq:\*
- 62: /cgn2\_6/ptodata/1/pna/US108A COMB.seq:\*
- 63: /cgn2\_6/ptodata/1/pna/US108B COMB.seq:\*
- 64: /cgn2\_6/ptodata/1/pna/US109A COMB.seq:\*
- 65: /cgn2\_6/ptodata/1/pna/US109B COMB.seq:\*
- 66: /cgn2\_6/ptodata/1/pna/US109C COMB.seq:\*
- 67: /cgn2\_6/ptodata/1/pna/US109D COMB.seq:\*
- 68: /cgn2\_6/ptodata/1/pna/US109E COMB.seq:\*
- 69: /cgn2\_6/ptodata/1/pna/US109F COMB.seq:\*
- 70: /cgn2\_6/ptodata/1/pna/US110A COMB.seq:\*
- 71: /cgn2\_6/ptodata/1/pna/US110B COMB.seq:\*
- 72: /cgn2\_6/ptodata/1/pna/US110C COMB.seq:\*
- 73: /cgn2\_6/ptodata/1/pna/US110D COMB.seq:\*
- 74: /cgn2\_6/ptodata/1/pna/US110E COMB.seq:\*
- 75: /cgn2\_6/ptodata/1/pna/US110F COMB.seq:\*
- 76: /cgn2\_6/ptodata/1/pna/US110G COMB.seq:\*
- 77: /cgn2\_6/ptodata/1/pna/US110H COMB.seq:\*
- 78: /cgn2\_6/ptodata/1/pna/US110I COMB.seq:\*
- 79: /cgn2\_6/ptodata/1/pna/US110J COMB.seq:\*
- 80: /cgn2\_6/ptodata/1/pna/US110K COMB.seq:\*
- 81: /cgn2\_6/ptodata/1/pna/US110L COMB.seq:\*
- 82: /cgn2\_6/ptodata/1/pna/US110M COMB.seq:\*
- 83: /cgn2\_6/ptodata/1/pna/US110N COMB.seq:\*
- 84: /cgn2\_6/ptodata/1/pna/US110O COMB.seq:\*
- 85: /cgn2\_6/ptodata/1/pna/US110P COMB.seq:\*
- 86: /cgn2\_6/ptodata/1/pna/US110Q COMB.seq:\*
- 87: /cgn2\_6/ptodata/1/pna/US110R COMB.seq:\*
- 88: /cgn2\_6/ptodata/1/pna/US110S COMB.seq:\*
- 89: /cgn2\_6/ptodata/1/pna/US110T COMB.seq:\*
- 90: /cgn2\_6/ptodata/1/pna/US110U COMB.seq:\*
- 91: /cgn2\_6/ptodata/1/pna/US110V COMB.seq:\*
- 92: /cgn2\_6/ptodata/1/pna/US110W COMB.seq:\*
- 93: /cgn2\_6/ptodata/1/pna/US110X COMB.seq:\*
- 94: /cgn2\_6/ptodata/1/pna/US110Y COMB.seq:\*
- 95: /cgn2\_6/ptodata/1/pna/US110Z COMB.seq:\*
- 96: /cgn2\_6/ptodata/1/pna/US110A COMB.seq:\*
- 97: /cgn2\_6/ptodata/1/pna/US110B COMB.seq:\*
- 98: /cgn2\_6/ptodata/1/pna/US110C COMB.seq:\*
- 99: /cgn2\_6/ptodata/1/pna/US110D COMB.seq:\*
- 100: /cgn2\_6/ptodata/1/pna/US110E COMB.seq:\*
- 101: /cgn2\_6/ptodata/1/pna/US110F COMB.seq:\*
- 102: /cgn2\_6/ptodata/1/pna/US110G COMB.seq:\*
- 103: /cgn2\_6/ptodata/1/pna/US110H COMB.seq:\*
- 104: /cgn2\_6/ptodata/1/pna/US110I COMB.seq:\*
- 105: /cgn2\_6/ptodata/1/pna/US110J COMB.seq:\*
- 106: /cgn2\_6/ptodata/1/pna/US110K COMB.seq:\*
- 107: /cgn2\_6/ptodata/1/pna/US110L COMB.seq:\*
- 108: /cgn2\_6/ptodata/1/pna/US110M COMB.seq:\*
- 109: /cgn2\_6/ptodata/1/pna/US110N COMB.seq:\*
- 110: /cgn2\_6/ptodata/1/pna/US110O COMB.seq:\*
- 111: /cgn2\_6/ptodata/1/pna/US110P COMB.seq:\*
- 112: /cgn2\_6/ptodata/1/pna/US110Q COMB.seq:\*
- 113: /cgn2\_6/ptodata/1/pna/US110R COMB.seq:\*
- 114: /cgn2\_6/ptodata/1/pna/US110S COMB.seq:\*
- 115: /cgn2\_6/ptodata/1/pna/US110T COMB.seq:\*
- 116: /cgn2\_6/ptodata/1/pna/US110U COMB.seq:\*

```

117: /cgn2_6/ptodata/1/pna/US6051 COMB seq.*
118: /cgn2_6/ptodata/1/pna/US6052 COMB seq.*
119: /cgn2_6/ptodata/1/pna/US6053 COMB seq.*
120: /cgn2_6/ptodata/1/pna/US6054 COMB seq.*
121: /cgn2_6/ptodata/1/pna/US6055 COMB seq.*
122: /cgn2_6/ptodata/1/pna/US6056 COMB seq.*
123: /cgn2_6/ptodata/1/pna/US6057 COMB seq.*
124: /cgn2_6/ptodata/1/pna/US6058 COMB seq.*
125: /cgn2_6/ptodata/1/pna/US6059 COMB seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		ID	Description
		Match	Length		
1	492	100.0	660	PCT-US03-26864-172	Sequence 172, App
2	492	100.0	660	PCT-US03-26864-590	Sequence 590, App
3	492	100.0	660	PCT-US03-27107-172	Sequence 172, App
4	492	100.0	660	PCT-US03-27107-590	Sequence 590, App
5	492	100.0	889	US-10-170-235-36232	Sequence 36232, A
6	492	100.0	889	US-60-452-680-2683	Sequence 2683, A
7	492	100.0	894	US-09-731-872-218	Sequence 218, App
8	492	100.0	894	US-09-876-997-218	Sequence 218, App
9	492	100.0	894	US-06-169-629-350	Sequence 350, App
10	492	100.0	894	US-60-187-470-350	Sequence 350, App
11	492	100.0	1086	US-09-649-163-8727	Sequence 8727, App
12	492	100.0	1086	US-09-699-997-8992	Sequence 8992, App
13	492	100.0	1086	US-09-736-119-1704	Sequence 1704, App
14	492	100.0	1089	PCT-US02-05109-107	Sequence 107, App
15	492	100.0	1089	US-09-810-173-107	Sequence 107, App
16	492	100.0	1104	US-09-649-163-6569	Sequence 6569, App
17	492	100.0	1104	US-09-652-913-7944	Sequence 7944, App
18	492	100.0	1178	US-09-758-472-1869	Sequence 1869, App
19	492	100.0	1178	US-10-235-926-1869	Sequence 1869, App
20	484	98.4	884	US-09-215-435-353	Sequence 353, App
21	484	98.4	884	US-09-978-360A-311	Sequence 311, App
22	484	97.6	884	US-60-069-957-186	Sequence 186, App
23	480	97.6	645	US-09-770-219-1	Sequence 1, Appli
24	470	95.5	1086	US-09-205-070-2956	Sequence 2956, App
25	470	95.5	1086	US-09-340-623-2956	Sequence 2956, App
26	470	95.5	1086	US-09-898-888-2956	Sequence 2956, App
27	470	95.5	1086	US-09-898-888A-2956	Sequence 2956, App
28	443.8	90.2	1091	US-09-359-322-2259	Sequence 2259, App
29	443.8	90.2	1091	US-09-359-322-2259	Sequence 2259, App
30	443.8	90.2	1091	US-09-359-322-2259	Sequence 2259, App
31	422.6	85.9	508	US-09-919-002-2259	Sequence 2259, App
32	417	84.8	575	US-09-540-212A-44786	Sequence 44786, A
33	394.2	80.1	542	US-09-540-229-99283	Sequence 99283, A
34	390.8	79.4	575	US-09-540-233D-106849	Sequence 106849, A
35	390.8	79.4	575	US-09-306-609-7116	Sequence 7116, App
36	390.8	79.4	575	US-09-522-251-7116	Sequence 7116, App
37	385.6	78.4	1036	US-10-025-600-7116	Sequence 7116, App
38	385.6	78.4	1036	PCT-US00-08983-59	Sequence 59, Appli
39	385.6	78.4	1036	PCT-US02-09105-132	Sequence 132, App
40	385.6	78.4	1036	PCT-US02-09135-93	Sequence 93, Appli
41	385.6	78.4	1036	PCT-US02-09257-163	Sequence 163, App
42	385.6	78.4	1036	PCT-US02-09322-162	Sequence 162, App
43	385.6	78.4	1036	US-09-950-082-335	Sequence 335, App
44	385.6	78.4	1036	US-10-105-299-335	Sequence 335, App
45	385.6	78.4	1036	US-10-472-964-162	Sequence 162, App
				US-10-472-965-163	Sequence 163, App

## ALIGNMENTS

RESULT 1  
PCT-US03-26864-172  
; Sequence 172, Application PC/TUS0326864  
; GENERAL INFORMATION:

```

; APPLICANT: FIVEPRIME THERAPEUTICS, INC.
; TITLE OF INVENTION: METHODS OF USE FOR NOVEL POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES
; FILE REFERENCE: 08940.0013-00304
; CURRENT APPLICATION NUMBER: PCT/US03/26864
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: 60/406,616
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,655
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,640
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,576
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,666
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,611
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,612
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/411,019
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,024
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,046
; PRIOR FILING DATE: 2002-09-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 731
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 172
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US03-26864-172

Query Match      100.0%; Score 492; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.4e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCAAGGCTTCCCAATGCCCCAG 60
Db      169  ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCAAGGCTTCCCAATGCCCCAG 228

QY      61  CTACCTCTTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACTGCTG 120
Db      229  CTACCTCTTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACTGCTG 288

QY      121  GGGTTGGCTGTGGGTCGGTTGAGGGCGGAGTGCTGGCATGTAGTGTCTCAGGTTCT 180
Db      289  GGGTTGGCTGTGGGTCGGTTGAGGGCGGAGTGCTGGCATGTAGTGTCTCAGGTTCT 348

QY      181  GGCAGGGCTGCAGGAAGGCTGTCACTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
Db      349  GGCAGGGCTGCAGGAAGGCTGTCACTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 408

QY      241  CTGCACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCCAGCTCA 300
Db      409  CTGCACAGCTCACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCCAGCTCA 468

QY      301  CCTGACACAGGGGTAGACCCCTTCACAGTGCGCCGCATGTGCTGCAGTGTGGGTGCTG 360
Db      469  CCTGACACAGGGGTAGACCCCTTCACAGTGCGCCGCATGTGCTGCAGTGTGGGTGCTG 528

QY      361  CTCAGCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCAAGGACCCCT 420
Db      529  CTCAGCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCAAGGACCCCT 588

QY      421  GGCCTGGGTTCATGCCAGCTCCAGCTGTGGGCCCTCGTTTCCCGAAGAGGGCTCGAGAC 480
Db      589  GGCCTGGGTTCATGCCAGCTCCAGCTGTGGGCCCTCGTTTCCCGAAGAGGGCTCGAGAC 648

QY      481  ACCGATCGTGA 492
Db      649  ACCGATCGTGA 660

```

RESULT 2  
PCT-US03-26864-590  
; Sequence 590, Application PC/TUS0326864  
; GENERAL INFORMATION:  
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.  
; TITLE OF INVENTION: METHODS OF USE FOR NOVEL POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES  
; FILE REFERENCE: 08940.0013-00304  
; CURRENT APPLICATION NUMBER: PCT/US03/26864  
; CURRENT FILING DATE: 2003-08-28  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,616  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,655  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,640  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,576  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,666  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,611  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,612  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/411,019  
; PRIOR FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: 60/411,024  
; PRIOR FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: 60/411,046  
; PRIOR FILING DATE: 2002-09-17  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 731  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 590  
; TYPE: DNA  
; LENGTH: 660  
; ORGANISM: Homo sapiens  
PCT-US03-26864-590

Query Match 100.0%; Score 492; DB 2; Length 660;  
Best Local Similarity 100.0%; Pred. No. 2.4e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 169 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 228  
QY 61 CTACCTCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 229 CTACCTCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 288  
QY 121 GGGTTGGCTCTGGGTGCGGTGGAGGGCGGCGAGTGTCCGATGTAGTGTCTCAGGTTCT 180  
Db 289 GGGTTGGCTCTGGGTGCGGTGGAGGGCGGCGAGTGTCCGATGTAGTGTCTCAGGTTCT 348  
QY 181 GGCAGGCTCGAGGAAGGCTGTGAGTGGCTGAGATTCTCAAGCGGGGTCGCCAGGC 240  
Db 349 GGCAGGCTCGAGGAAGGCTGTGAGTGGCTGAGATTCTCAAGCGGGGTCGCCAGGC 408  
QY 241 CTGCACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 300  
Db 409 CTGCACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 468  
QY 301 CTGCACAGGCTAGACCCCTTCAAGTGGTGTGTTACCAACCCCGAGGACACCCCT 420  
Db 469 CTGCACAGGCTAGACCCCTTCAAGTGGTGTGTTACCAACCCCGAGGACACCCCT 588  
QY 421 GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTTCCCGAAGAGGGCTCGAGAC 480

Db 589 GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTCCGAGAGAGGGCTCGAGAC 648  
QY 481 ACCCGATCGTGA 492  
Db 649 ACCCGATCGTGA 660

RESULT 3  
PCT-US03-27107-172  
; Sequence 172, Application PC/TUS0327107  
; GENERAL INFORMATION:  
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.  
; TITLE OF INVENTION: NOVEL HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES  
; FILE REFERENCE: 08940.0011-00304  
; CURRENT APPLICATION NUMBER: PCT/US03/27107  
; CURRENT FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: 60/406,616  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,655  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,640  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,576  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,666  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,611  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,612  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/411,019  
; PRIOR FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: 60/411,024  
; PRIOR FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: 60/411,046  
; PRIOR FILING DATE: 2002-09-17  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 731  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 172  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
PCT-US03-27107-172

Query Match 100.0%; Score 492; DB 2; Length 660;  
Best Local Similarity 100.0%; Pred. No. 2.4e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 169 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 228  
QY 61 CTACCTCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 229 CTACCTCTCTGATACCTTGGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 288  
QY 121 GGGTTGGCTCTGGGTGCGGTGGAGGGCGGCGAGTGTCCGATGTAGTGTCTCAGGTTCT 180  
Db 289 GGGTTGGCTCTGGGTGCGGTGGAGGGCGGCGAGTGTCCGATGTAGTGTCTCAGGTTCT 348  
QY 181 GGCAGGCTCGAGGAAGGCTGTGAGTGGCTGAGATTCTCAAGCGGGGTCGCCAGGC 240  
Db 349 GGCAGGCTCGAGGAAGGCTGTGAGTGGCTGAGATTCTCAAGCGGGGTCGCCAGGC 408  
QY 241 CTGCACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 300  
Db 409 CTGCACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 468  
QY 301 CTGCACAGGCTAGACCCCTTCAAGTGGTGTGTTACCAACCCCGAGGACACCCCT 360  
Db 469 CTGCACAGGCTAGACCCCTTCAAGTGGTGTGTTACCAACCCCGAGGACACCCCT 528

Qy	361	CTCAGCCGGGACCCCTCGAACCCCAATCAGTGTGGTTACCAACCCCGAGGAGCACCCCTT	420
Db	529	CTCAGCCGGGACCCCTCGAACCCCAATCAGTGTGGTTACCAACCCCGAGGAGCACCCCTT	588
Qy	421	GGCTTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCGAAAGAGGGTCGAGAC	480
Db	589	GGCTTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCGAAAGAGGGTCGAGAC	648
Qy	481	ACCCGATCGTGA	492
Db	649	ACCCGATCGTGA	660

```

RESULT 4
PCT-US03-27107-590
; Sequence 590, Application PC/TUS0327107
; GENERAL INFORMATION:
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.
; TITLE OF INVENTION: NOVEL HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES
; FILE REFERENCE: 08940.0011-00304
; CURRENT APPLICATION NUMBER: PCT/US03/27107
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: 60/406,616
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,655
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,640
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,576
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,666
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,611
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,612
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/411,019
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,024
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,046
; PRIOR FILING DATE: 2002-09-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 731
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 590
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US03-27107-590

```

Query Match	100.0%	Score 492;	DB 2;	Length 660;
Best Local Similarity	100.0%;	Pred. No. 2.4e-110;		
Matches 492;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAGACACTACCGGAAGCTGGCTCTGTAGAGCTCCACGCGCTTCCCAATGCCACG	60	
Db	169	ATGAGACACTACCGGAAGCTGGCTCTGTAGAGCTCCACGCGCTTCCCAATGCCACG	228	
QY	61	CTACCTCGTATACCCCTTGATAGATGCGGGTCCGAGATGGCAGCAAAATTGCAAACTGTGTG	120	
Db	229	CTACCTCGTATACCCCTTGATAGATGCGGGTCCGAGATGGCAGCAAAATTGCAAACTGTGTG	288	
QY	121	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTTCAGGTTCT	180	
Db	289	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTTCAGGTTCT	348	
QY	181	GGCAGGCTCCAGGAAGGCTGTCACTGGCTGAGATTGTCAAGCGGGGGTCCCAAGC	240	
Db	349	GGCAGGCTCCAGGAAGGCTGTCACTGGCTGAGATTGTCAAGCGGGGGTCCCAAGC	408	
QY	241	CTGCACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCCAAGCCTCA	300	

Db	409	CTGACACGAGCTCACCAAGCTAGTTTCTTTCAGACTGAGGACAGCTGGGTCCCAGGCTCA	468
Qy	301	CCTGACACAGGGCTAGACCCCTTCACAGTGGCCGCCCATGTGCTTCAGTGTGGGTGCTG	360
Db	469	CCTGACACAGGGCTAGACCCCTTCACAGTGGCCGCCCATGTGCTTCAGTGTGGGTGCTG	528
Qy	361	CTCAGCCGGGACCCCTTGAGCCCAATGAGTGTGTTACCAACCCCGCAGGAGCACCCCTCT	420
Db	529	CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGTTACCAACCCCGCAGGAGCACCCCTCT	588
Qy	421	GGCTGTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCTCCGAAGAGGGCTCGAGAC	480
Db	589	GGCTGTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCTCCGAAGAGGGCTCGAGAC	648
Qy	481	ACCCGATCGTGA	492
Db	649	ACCCGATCGTGA	660

```

RESULT 5
US-10-170-235-36232
; Sequence 36232, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC
; FILE OF INVENTION: TRANSCRIPTS, FOR DET
; TITLE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 36232
; LENGTH: 889
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-36232

```

Query Match	100.0%;	Score 492;	DB 48;	Length 889;
Best Local Similarity	100.0%;	Pred. No. 2.5e-110;		
Matches 492;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCACAGCGCTTCCCAATGCCCGAG	60	
Db	79	ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCGAG	138	
QY	61	CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGACGCAAAATTCGCAACCTGCTG	120	
Db	139	CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGACGCAAAATTCGCAACCTGCTG	198	
QY	121	GGGTGGCTCTGGTTCGGTTGGAGGGCGGAGTGCTCGGATGTAGTGTCTCAGAGTCT	180	
Db	199	GGGTGGCTCTGGTTCGGTTGGAGGGCGGAGTGCTCGGATGTAGTGTCTCAGAGTCT	258	
QY	181	GGCAGGGCTCGAGAAAGGCTGTCACTGCGTCAAGTTCGAGCGCGGGTCCCAAGC	240	
Db	259	GGCAGGGCTCGAGAAAGGCTGTCACTGCGTCAAGTTCGAGCGCGGGTCCCAAGC	318	
QY	241	CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAAGCTCA	300	
Db	319	CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAAGCTCA	378	
QY	301	CCTGACACAGGGCTAGACCCCTTCACAGTGCAGCGCGCCATGTGCTGAGTGTGGGTGCTG	360	
Db	379	CCTGACACAGGGCTAGACCCCTTCACAGTGCAGCGCGCCATGTGCTGAGTGTGGGTGCTG	438	
QY	361	CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT	420	
Db	439	CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT	498	
QY	421	GGCTGGGTTCATGCCAGCTCCAGCTGAGCGCCCTCGTTCCGAGAAAGGGCTCGAGAC	480	
Db	499	GGCTGGGTTCATGCCAGCTCCAGCTGAGCGCCCTCGTTCCGAGAAAGGGCTCGAGAC	558	



QY 481 ACCCGATCGTGA 492  
Db 559 ACCCGATCGTGA 570

## RESULT 6

US-60-452-680-2683  
; Sequence 2683, Application US/60452680  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: GRUPE, Andrew  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001450  
; CURRENT APPLICATION NUMBER: US/60/452,680  
; CURRENT FILING DATE: 2003-03-07  
; NUMBER OF SEQ ID NOS: 116213  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 2683  
; LENGTH: 889  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-452-680-2683

Query Match 100.0%; Score 492; DB 111; Length 889;  
Best Local Similarity 100.0%; Pred. No. 2.5e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 79 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 138  
QY 61 CTACCTCTCTGATACCTTGTAGATCGGGTCCGAGATGGCAAAATTCGCAACCTGCTG 120  
Db 139 CTACCTCTCTGATACCTTGTAGATCGGGTCCGAGATGGCAAAATTCGCAACCTGCTG 198  
QY 121 GGGTTGGCTCTGGGTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
Db 199 GGGTTGGCTCTGGGTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 258  
QY 181 GGCAGGCTCGAGAAAGCTGTGAGTGGCTGAGATGTCAGGGCGGGTCCAGGC 240  
Db 259 GGCAGGCTCGAGAAAGCTGTGAGTGGCTGAGATGTCAGGGCGGGTCCAGGC 318  
QY 241 CTCACACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCCTCA 300  
Db 319 CTCACACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCCTCA 378  
QY 301 CCTGACACAGGCTAGACCCCTTCAAGTGGCGGCCCATGTGCTGAGTGTGGGTGCTG 360  
Db 379 CCTGACACAGGCTAGACCCCTTCAAGTGGCGGCCCATGTGCTGAGTGTGGGTGCTG 438  
QY 361 CTCAGCGGAGCCCTTGGACCCCAATGAGTGTGGTGTACCAACCCCGAGGACCCCTT 420  
Db 439 CTCAGCGGAGCCCTTGGACCCCAATGAGTGTGGTGTACCAACCCCGAGGACCCCTT 498  
QY 421 GGCCTGGGTTCATGCCAGCTCCAGCTGGGCCCTCGTTCCGAAAGAGGGCTCGAGAC 480  
Db 499 GGCCTGGGTTCATGCCAGCTCCAGCTGGGCCCTCGTTCCGAAAGAGGGCTCGAGAC 558  
QY 481 ACCCGATCGTGA 492  
Db 559 ACCCGATCGTGA 570

## RESULT 7

US-09-731-872-218  
; Sequence 218, Application US/09731872  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS

; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731.872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 218  
; LENGTH: 894  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 78..566  
US-09-731-872-218

Query Match 100.0%; Score 492; DB 32; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.5e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 78 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 137  
QY 61 CTACCTCTCTGATACCTTGTAGATCGGGTCCGAGATGGCAAAATTCGCAACCTGCTG 120  
Db 138 CTACCTCTCTGATACCTTGTAGATCGGGTCCGAGATGGCAAAATTCGCAACCTGCTG 197  
QY 121 GGGTTGGCTCTGGGTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
Db 198 GGGTTGGCTCTGGGTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 257  
QY 181 GGCAGGCTCGAGAAAGCTGTGAGTGGCTGAGATGTCAGGGCGGGTCCAGGC 240  
Db 258 GGCAGGCTCGAGAAAGCTGTGAGTGGCTGAGATGTCAGGGCGGGTCCAGGC 317  
QY 241 CTCACACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCCTCA 300  
Db 318 CTCACACAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCCTCA 377  
QY 301 CCTGACACAGGCTAGACCCCTTCAAGTGGCGGCCCATGTGCTGAGTGTGGGTGCTG 360  
Db 378 CCTGACACAGGCTAGACCCCTTCAAGTGGCGGCCCATGTGCTGAGTGTGGGTGCTG 437  
QY 361 CTCAGCGGAGCCCTTGGACCCCAATGAGTGTGGTGTACCAACCCCGAGGACCCCTT 420  
Db 438 CTCAGCGGAGCCCTTGGACCCCAATGAGTGTGGTGTACCAACCCCGAGGACCCCTT 497  
QY 421 GGCCTGGGTTCATGCCAGCTCCAGCTGGGCCCTCGTTCCGAAAGAGGGCTCGAGAC 480  
Db 498 GGCCTGGGTTCATGCCAGCTCCAGCTGGGCCCTCGTTCCGAAAGAGGGCTCGAGAC 557  
QY 481 ACCCGATCGTGA 492  
Db 558 ACCCGATCGTGA 569

## RESULT 8

US-09-876-997-218  
; Sequence 218, Application US/09876997  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US4.CIP  
; CURRENT APPLICATION NUMBER: US/09/876,997  
; CURRENT FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 09/731.872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/187,470

;; PRIOR FILING DATE: 2000-03-06  
;; PRIOR APPLICATION NUMBER: US 60/169,629  
;; PRIOR FILING DATE: 1999-12-08  
;; NUMBER OF SEQ ID NOS: 482  
;; SOFTWARE: Patent.pm  
;; SEQ ID NO 218  
;; LENGTH: 894  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 78..566  
US-09-876-997-218

Query Match 100.0%; Score 492; DB 36; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.5e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 78 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 137  
QY 61 CTACCTCTGATACCTTGAGATGCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 120  
Db 138 CTACCTCTGATACCTTGAGATGCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 197  
QY 121 GGGTTGGCTCTGGGTGGGTGGAGGGCGGAGTGGCTCGGCATGTAGTGTCTCAGGTTCT 180  
Db 198 GGGTTGGCTCTGGGTGGGTGGAGGGCGGAGTGGCTCGGCATGTAGTGTCTCAGGTTCT 257  
QY 181 GGCAGGGCTGCAGGAAAGCTGTGAGCTGGCTGAGATGTCTAAGGGGGGGTCCAGGC 240  
Db 258 GGCAGGGCTGCAGGAAAGCTGTGAGCTGGCTGAGATGTCTAAGGGGGGGTCCAGGC 317  
QY 241 CTGACACAGCTACCAAGCTACCTTTCCTTTCAGCTGAGAGACAGCTGGGTCCAGCCTCA 300  
Db 318 CTGACACAGCTACCAAGCTACCTTTCCTTTCAGCTGAGAGACAGCTGGGTCCAGCCTCA 377  
QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGCAGCGCCGATGTGCTGAGTGGGTGCTG 360  
Db 378 CCTGACACAGGGCTAGACCCCTTACAGTGCAGCGCCGATGTGCTGAGTGGGTGCTG 437  
QY 361 CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420  
Db 438 CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 497  
QY 421 GGCTGGGTTCATGCCAGCTCAGCTGGGCCCTGCTTCCGAAAGGGGCTCGAGAC 480  
Db 498 GGCTGGGTTCATGCCAGCTCAGCTGGGCCCTGCTTCCGAAAGGGGCTCGAGAC 557  
QY 481 ACCCGATCGTGA 492  
Db 558 ACCCGATCGTGA 569

RESULT 9  
US-60-169-629-350  
;; Sequence 350, Application US/60169629  
;; GENERAL INFORMATION:  
;; APPLICANT: Dumas Milne Edwards, J.B.  
;; APPLICANT: Bougueleret, L.  
;; APPLICANT: Jobert, S.  
;; TITLE OF INVENTION: cDNAs for Secreted Proteins  
;; FILE REFERENCE: GENSET.071PF  
;; CURRENT APPLICATION NUMBER: US/60/169,629  
;; CURRENT FILING DATE: 1999-12-08  
;; NUMBER OF SEQ ID NOS: 715  
;; SOFTWARE: Patent.pm  
;; SEQ ID NO 350  
;; LENGTH: 894  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:

;; NAME/KEY: CDS  
;; LOCATION: 78..566  
;; FEATURE:  
;; NAME/KEY: sig\_peptide  
;; LOCATION: 78..233  
;; OTHER INFORMATION: Von Heijne matrix  
;; OTHER INFORMATION: score 4.7  
;; OTHER INFORMATION: seq LGLALGRLEGGSA/RH  
;; FEATURE:  
;; NAME/KEY: polyA\_signal  
;; LOCATION: 858..863  
;; FEATURE:  
;; NAME/KEY: polyA\_site  
;; LOCATION: 878..894  
US-60-169-629-350

Query Match 100.0%; Score 492; DB 81; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.5e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 78 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 137  
QY 61 CTACCTCTGATACCTTTGAGATGCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 120  
Db 138 CTACCTCTGATACCTTTGAGATGCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 197  
QY 121 GGGTTGGCTCTGGGTGGGTGGAGGGCGGAGTGGCTCGGCATGTAGTGTCTCAGGTTCT 180  
Db 198 GGGTTGGCTCTGGGTGGGTGGAGGGCGGAGTGGCTCGGCATGTAGTGTCTCAGGTTCT 257  
QY 181 GGCAGGGCTGCAGGAAAGCTGTGAGCTGGCTGAGATGTCTAAGGGGGGGTCCAGGC 240  
Db 258 GGCAGGGCTGCAGGAAAGCTGTGAGCTGGCTGAGATGTCTAAGGGGGGGTCCAGGC 317  
QY 241 CTGACACAGCTACCAAGCTACCTTTCCTTTCAGCTGAGAGACAGCTGGGTCCAGCCTCA 300  
Db 318 CTGACACAGCTACCAAGCTACCTTTCCTTTCAGCTGAGAGACAGCTGGGTCCAGCCTCA 377  
QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGCAGCGCCGATGTGCTGAGTGGGTGCTG 360  
Db 378 CCTGACACAGGGCTAGACCCCTTACAGTGCAGCGCCGATGTGCTGAGTGGGTGCTG 437  
QY 361 CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420  
Db 438 CTCAGCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 497  
QY 421 GGCTGGGTTCATGCCAGCTCAGCTGGGCCCTGCTTCCGAAAGGGGCTCGAGAC 480  
Db 498 GGCTGGGTTCATGCCAGCTCAGCTGGGCCCTGCTTCCGAAAGGGGCTCGAGAC 557  
QY 481 ACCCGATCGTGA 492  
Db 558 ACCCGATCGTGA 569

RESULT 10  
US-60-187-470-350  
;; Sequence 350, Application US/60187470  
;; GENERAL INFORMATION:  
;; APPLICANT: Dumas Milne Edwards, J.B.  
;; APPLICANT: Bougueleret, L.  
;; APPLICANT: Jobert, S.  
;; TITLE OF INVENTION: cDNAs for Secreted Proteins  
;; FILE REFERENCE: 78.US2.PRO  
;; CURRENT APPLICATION NUMBER: US/60/187,470  
;; CURRENT FILING DATE: 2000-03-06  
;; NUMBER OF SEQ ID NOS: 715  
;; SOFTWARE: Patent.pm  
;; SEQ ID NO 350  
;; LENGTH: 894  
;; TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 78..566  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 78..233  
OTHER INFORMATION: Von Heijne matrix  
OTHER INFORMATION: score 4.7  
OTHER INFORMATION: seq LGLALGRLEGGSA/RH  
FEATURE:  
NAME/KEY: polyA signal  
LOCATION: 858..563  
FEATURE:  
NAME/KEY: polyA\_site  
LOCATION: 878..894  
US-60-187-470-350

Query Match 100.0%; Score 492; DB 83; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.5e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
DB 78 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 137  
QY 61 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGCGCAAAATTCGCAACCTGCTG 120  
DB 138 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGCGCAAAATTCGCAACCTGCTG 197  
QY 121 GGGTTGGCTCTGGGTCTGGTTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
DB 198 GGGTTGGCTCTGGGTCTGGTTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 257  
QY 181 GGCAGGGCTCGAGAAAGGTGTCTGAGTGGCTGGAGATGTCGAGTGTCTCAGGTTCT 240  
DB 258 GGCAGGGCTCGAGAAAGGTGTCTGAGTGGCTGGAGATGTCGAGTGTCTCAGGTTCT 317  
QY 241 CTGACACAGCTACCAAGCTACGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 300  
DB 318 CTGACACAGCTACCAAGCTACGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 377  
QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGGCGGCCCATGTGCTGAGTGTGGGTGCTG 360  
DB 378 CCTGACACAGGGCTAGACCCCTTACAGTGGCGGCCCATGTGCTGAGTGTGGGTGCTG 437  
QY 361 CTGACCGGAGACCCCTTGGACCCCATAGTGTGGTTTACCAACCCCGAGGACACCCCT 420  
DB 438 CTGACCGGAGACCCCTTGGACCCCATAGTGTGGTTTACCAACCCCGAGGACACCCCT 497  
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCTGTTCCCGAAGAGGGCTCGAGAC 480  
DB 498 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCTGTTCCCGAAGAGGGCTCGAGAC 557  
QY 481 ACCCGATCGTGA 492  
DB 558 ACCCGATCGTGA 569

RESULT 11  
US-09-649-163-8727  
Sequence 8727, Application US/09649163  
GENERAL INFORMATION:  
APPLICANT: Holtzman, Douglas A.  
APPLICANT: Galvin, Katherine A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kingsbury, Gillian A.  
APPLICANT: Welch, Nadine S.  
APPLICANT: McCarthy, Sean A.  
APPLICANT: Williamson, Mark  
APPLICANT: Richardson, Jennifer  
APPLICANT: MacBeth, Kyle J.  
APPLICANT: Fraser, Christopher C.

APPLICANT: Villevall, Jean-Luc M.G.  
APPLICANT: Goodearl, Andrew D.J.  
APPLICANT: Silos-Santiago, Inmaculada  
APPLICANT: White, David  
APPLICANT: Pan, Yang  
APPLICANT: Busfield, Samantha J.  
APPLICANT: Deeds, James  
APPLICANT: Lee, John  
APPLICANT: Shyjan, Andrew W.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: 1600.1164-001  
CURRENT APPLICATION NUMBER: US/09/649,163  
CURRENT FILING DATE: 2000-08-25  
PRIOR APPLICATION NUMBER: 60/150,608  
PRIOR FILING DATE: 1999-08-25  
NUMBER OF SEQ ID NOS: 10535  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8727  
LENGTH: 1086  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-649-163-8727

Query Match 100.0%; Score 492; DB 28; Length 1086;  
Best Local Similarity 100.0%; Pred. No. 2.6e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
DB 275 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 334  
QY 61 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGCGCAAAATTCGCAACCTGCTG 120  
DB 335 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGCGCAAAATTCGCAACCTGCTG 394  
QY 121 GGGTTGGCTCTGGGTCTGGTTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
DB 395 GGGTTGGCTCTGGGTCTGGTTGGAGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 454  
QY 181 GGCAGGGCTCGAGAAAGGTGTCTGAGTGGCTGGAGATGTCGAGTGTGGGTGCTG 240  
DB 455 GGCAGGGCTCGAGAAAGGTGTCTGAGTGGCTGGAGATGTCGAGTGTGGGTGCTG 514  
QY 241 CTGACACAGCTACCAAGCTACGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 300  
DB 515 CTGACACAGCTACCAAGCTACGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 574  
QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGGCGGCCCATGTGCTGAGTGTGGGTGCTG 360  
DB 575 CCTGACACAGGGCTAGACCCCTTACAGTGGCGGCCCATGTGCTGAGTGTGGGTGCTG 634  
QY 361 CTGACCGGAGACCCCTTGGACCCCATAGTGTGGTTTACCAACCCCGAGGACACCCCT 420  
DB 635 CTGACCGGAGACCCCTTGGACCCCATAGTGTGGTTTACCAACCCCGAGGACACCCCT 694  
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCTGTTCCCGAAGAGGGCTCGAGAC 480  
DB 695 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCTGTTCCCGAAGAGGGCTCGAGAC 754  
QY 481 ACCCGATCGTGA 492  
DB 755 ACCCGATCGTGA 766

RESULT 12  
US-09-699-997-8992  
Sequence 8992, Application US/09699997  
GENERAL INFORMATION:  
APPLICANT: Curtis, Rory A.J.  
APPLICANT: Silos-Santiago, Inmaculada  
APPLICANT: DiStefano, Peter  
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

```
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2020-001
; CURRENT APPLICATION NUMBER: US/09/699,997
; CURRENT FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 60/162,359
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 12714
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8992
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-699-997-8992

Query Match      100.0%; Score 492; DB 30; Length 1086;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
DB 275 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 334
QY 61 CTACCTCTGATACCTTCCGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
DB 335 CTACCTCTGATACCTTCCGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 394
QY 121 GGGTTGGCTCTGGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTC 180
DB 395 GGGTTGGCTCTGGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTC 454
QY 181 GGCAGGGCTCGAGGAAAGGCTGTGAGTGGCTCGAGTGGCTCGAGTGGCTCGAGTGGCT 240
DB 455 GGCAGGGCTCGAGGAAAGGCTGTGAGTGGCTCGAGTGGCTCGAGTGGCTCGAGTGGCT 514
QY 241 CTCACCAAGCTACCAAGCTACCTTCCGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 300
DB 515 CTCACCAAGCTACCAAGCTACCTTCCGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 574
QY 301 CCTGACACAGGGCTAGACCCCTTCAGTGGCTCGAGTGGCTCGAGTGGCTCGAGTGGCT 360
DB 575 CCTGACACAGGGCTAGACCCCTTCAGTGGCTCGAGTGGCTCGAGTGGCTCGAGTGGCT 634
QY 361 CTCAGCGGGGACCCCTTGAGACCCCAATGAGTGGTGTACCAACCCCGAGGAGCACCCCT 420
DB 635 CTCAGCGGGGACCCCTTGAGACCCCAATGAGTGGTGTACCAACCCCGAGGAGCACCCCT 694
QY 421 GGCTTGGGTTCATGCCAGCTCCAGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
DB 695 GGCTTGGGTTCATGCCAGCTCCAGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 754
QY 481 ACCCGATCGTGA 492
DB 755 ACCCGATCGTGA 766

RESULT 13
US-09-736-119-1704
; Sequence 1704, Application US/09736119
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Villevall, Jean-Luc M.G.
; APPLICANT: Fraser, Christopher C.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2047-001
; CURRENT APPLICATION NUMBER: US/09/736,119
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/170,468
; PRIOR FILING DATE: 1999-12-13
; NUMBER OF SEQ ID NOS: 2118
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1704
; LENGTH: 1086
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-119-1704

Query Match      100.0%; Score 492; DB 32; Length 1086;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
DB 275 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 334
QY 61 CTACCTCTGATACCTTCCGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
DB 335 CTACCTCTGATACCTTCCGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 394
QY 121 GGGTTGGCTCTGGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTC 180
DB 395 GGGTTGGCTCTGGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTC 454
QY 181 GGCAGGGCTCGAGGAAAGGCTGTGAGTGGCTCGAGTGGCTCGAGTGGCTCGAGTGGCT 240
DB 455 GGCAGGGCTCGAGGAAAGGCTGTGAGTGGCTCGAGTGGCTCGAGTGGCTCGAGTGGCT 514
QY 241 CTCACCAAGCTACCAAGCTACCTTCCGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 300
DB 515 CTCACCAAGCTACCAAGCTACCTTCCGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 574
QY 301 CCTGACACAGGGCTAGACCCCTTCAGTGGCTCGAGTGGCTCGAGTGGCTCGAGTGGCT 360
DB 575 CCTGACACAGGGCTAGACCCCTTCAGTGGCTCGAGTGGCTCGAGTGGCTCGAGTGGCT 634
QY 361 CTCAGCGGGGACCCCTTGAGACCCCAATGAGTGGTGTACCAACCCCGAGGAGCACCCCT 420
DB 635 CTCAGCGGGGACCCCTTGAGACCCCAATGAGTGGTGTACCAACCCCGAGGAGCACCCCT 694
QY 421 GGCTTGGGTTCATGCCAGCTCCAGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
DB 695 GGCTTGGGTTCATGCCAGCTCCAGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 754
QY 481 ACCCGATCGTGA 492
DB 755 ACCCGATCGTGA 766

RESULT 14
PCT-US02-05109-107
; Sequence 107, Application PC/TUS0205109
; GENERAL INFORMATION:
; APPLICANT: HySeq Inc.
; APPLICANT: Tang, Tom Y.
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Ma, Yuning
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-094-061
; CURRENT APPLICATION NUMBER: PCT/US02/05109
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 09/810,173
```

; PRIOR FILING DATE: 2001-03-15  
; NUMBER OF SEQ ID NOS: 1052  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 107  
; LENGTH: 1089  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (275)..(766)  
PCT-US02-05109-107

Query Match 100.0%; Score 492; DB 2; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 2.6e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCCTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 275 ATGGAGCCTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 334  
QY 61 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 335 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 394  
QY 121 GGGTTGGCTCTGGGTGCGTTGGAGGCGCGAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
Db 395 GGGTTGGCTCTGGGTGCGTTGGAGGCGCGAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 454  
QY 181 GGCAGGCTCCAGGAAGCTGTCAGCTGCGTGGAGATGTCAGCGCGGGTCCAGGC 240  
Db 455 GGCAGGCTCCAGGAAGCTGTCAGCTGCGTGGAGATGTCAGCGCGGGTCCAGGC 514  
QY 241 CTGACACAGCTCACCAGCTACGTTTCTTCCAGTGGAGCAGCTGGGTCCAGGCTCA 300  
Db 515 CTGACACAGCTCACCAGCTACGTTTCTTCCAGTGGAGCAGCTGGGTCCAGGCTCA 574  
QY 301 CCTGACACAGGCTAGACCCCTTCCAGTGGCGGCCATGTGCTGCAAGTGTGGGTGCTG 360  
Db 575 CCTGACACAGGCTAGACCCCTTCCAGTGGCGGCCATGTGCTGCAAGTGTGGGTGCTG 634  
QY 361 CTGACGCGGACCCCTGGAGCCCAATGATGTGGTGTACCAACCCCGAGGACACCCCT 420  
Db 635 CTGACGCGGACCCCTGGAGCCCAATGATGTGGTGTACCAACCCCGAGGACACCCCT 694  
QY 421 GGCCTGGTTCATGCCAGCTCCAGTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480  
Db 695 GGCCTGGTTCATGCCAGCTCCAGTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 754  
QY 481 ACCGATCGTGA 492  
Db 755 ACCGATCGTGA 766

## RESULT 15

US-09-810-173-107  
; Sequence 107, Application US/09810173  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyun  
; APPLICANT: Zhang, Jie  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Ma, Yuning  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Yamazaki, Victoria  
; APPLICANT: Wang, Dunrui

; APPLICANT: Ghosh, Malabika J.  
; APPLICANT: Dimanac, Radoje T.  
; TITLE OF INVENTION: Novel Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 804  
; CURRENT APPLICATION NUMBER: US/09/810,173  
; CURRENT FILING DATE: 2001-03-15  
; NUMBER OF SEQ ID NOS: 526  
; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 107  
; LENGTH: 1089  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (275)..(763)  
US-09-810-173-107

Query Match 100.0%; Score 492; DB 34; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 2.6e-110;  
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCCTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 275 ATGGAGCCTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 334  
QY 61 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 335 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 394  
QY 121 GGGTTGGCTCTGGGTGCGTTGGAGGCGCGAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
Db 395 GGGTTGGCTCTGGGTGCGTTGGAGGCGCGAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 454  
QY 181 GGCAGGCTCCAGGAAGCTGTCAGCTGCGTGGAGATGTCAGCGCGGGTCCAGGC 240  
Db 455 GGCAGGCTCCAGGAAGCTGTCAGCTGCGTGGAGATGTCAGCGCGGGTCCAGGC 514  
QY 241 CTGACACAGCTCACCAGCTACGTTTCTTCCAGTGGAGCAGCTGGGTCCAGGCTCA 300  
Db 515 CTGACACAGCTCACCAGCTACGTTTCTTCCAGTGGAGCAGCTGGGTCCAGGCTCA 574  
QY 301 CCTGACACAGGCTAGACCCCTTCCAGTGGCGGCCATGTGCTGCAAGTGTGGGTGCTG 360  
Db 575 CCTGACACAGGCTAGACCCCTTCCAGTGGCGGCCATGTGCTGCAAGTGTGGGTGCTG 634  
QY 361 CTGACGCGGACCCCTGGAGCCCAATGATGTGGTGTACCAACCCCGAGGACACCCCT 420  
Db 635 CTGACGCGGACCCCTGGAGCCCAATGATGTGGTGTACCAACCCCGAGGACACCCCT 694  
QY 421 GGCCTGGTTCATGCCAGCTCCAGTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480  
Db 695 GGCCTGGTTCATGCCAGCTCCAGTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 754  
QY 481 ACCGATCGTGA 492  
Db 755 ACCGATCGTGA 766

Search completed: January 31, 2005, 18:58:39  
Job time : 2955 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 12:16:25 ; Search time 85 Seconds  
(without alignments)  
4114.213 Million cell updates/sec

Title: US-10-057-813-13

Perfect score: 492

Sequence: 1 atggagcactaccggaagc.....ctcgagacaccgatcgta 492

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	374.4	76.1	461	US-09-621-976-217	Sequence 217, App
2	236	48.0	303	US-09-513-999C-3603	Sequence 3603, Ap
3	74.6	15.2	1941	US-09-620-312D-5	Sequence 5, Appli
4	74.6	15.2	2375	US-09-976-594-92	Sequence 92, Appl
C 5	46	9.3	356	US-09-270-767-8210	Sequence 8210, Ap
C 6	46	9.3	356	US-09-270-767-23492	Sequence 23492, A
C 7	37.2	7.6	46819	US-09-453-702B-72	Sequence 72, Appl
8	35.6	7.2	22306	US-09-453-702B-251	Sequence 251, App
9	35.4	7.2	4257	US-08-690-473-1	Sequence 1, Appli
10	35.4	7.2	4257	US-09-259-821A-1	Sequence 1, Appli
11	35.4	7.2	4257	US-08-843-659-1	Sequence 1, Appli
12	35.4	7.2	4257	US-09-825-288A-1	Sequence 1, Appli
C 13	35.4	7.2	12001	US-08-458-568A-11	Sequence 11, Appl
C 14	35.2	7.2	4403765	US-09-103-840A-2	Sequence 2, Appli
15	34.8	7.1	1320	US-09-727-238-1	Sequence 1, Appli
C 16	34.8	7.1	2074	US-09-690-942-7	Sequence 7, Appli
17	34.6	7.0	531	US-09-252-991A-12044	Sequence 12044, A
C 18	34.4	7.0	1476	US-09-489-039A-5063	Sequence 5063, Ap
19	33	6.7	4403765	US-09-103-840A-2	Sequence 2, Appli
20	33	6.7	4411529	US-09-103-840A-1	Sequence 1, Appli
C 21	32.6	6.6	378	5274075-3	Patent No. 5274075
22	32.6	6.6	2430	US-09-252-991A-12709	Sequence 12709, A
C 23	32.6	6.6	2607	US-09-252-991A-13249	Sequence 13249, A
C 24	32.4	6.6	950	US-09-270-767-11073	Sequence 11073, A
C 25	31.8	6.5	1950	US-09-252-991A-11086	Sequence 11086, A
C 26	31.8	6.5	2127	US-09-252-991A-8192	Sequence 8192, Ap
27	31.8	6.5	2874	US-09-252-991A-8112	Sequence 8112, Ap

ALIGNMENTS

RESULT 1

US-09-621-976-217

; Sequence 217, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 217

; LENGTH: 461

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 78..461

; NAME/KEY: sig\_peptide

; LOCATION: 78..233

; OTHER INFORMATION: Von Heijne matrix

; OTHER INFORMATION: score 4.6999980926514

; OTHER INFORMATION: seq LGALGRLEGSSA/RH

US-09-621-976-217

Query Match 76.1%; Score 374.4; DB 4; Length 461;

Best Local Similarity 98.4%; Pred. No. 7.4e-98;

Matches 376; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY	1	ATGAGCCTACCGGAAAGCTGGCTCTGTAGAGTCCAGCGCTTCCCAATGCCCGAG 60
DB	78	ATGAGCCTACCGGAAAGCTGGCTCTGTAGAGTCCAGCGCTTCCCAATGCCCGAG 137
QY	61	CTACCTCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
DB	138	CTACCTCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 197
QY	121	GGTGTGGCTCTGGTCCGTTGGAGCGCGCAGTCTCGGCATGTAGTGTCTCAGATTCT 180
DB	198	GGTGTGGCTCTGGTCCGTTGGAGCGCGCAGTCTCGGCATGTAGTGTCTCAGATTCT 257
QY	181	GGCAGGGCTGCAGGAAGCTGTGAGTCTCGCTGAGATGTCAAGCGGGGGTCCCGAGC 240
DB	258	GGCAGGGCTGCAGGAAGCTGTGAGTCTCGCTGAGATGTCAAGCGGGGGTCCCGAGC 317
QY	241	CTGCACCAAGCTCACCAGCTAGGTTTCTTCCAGCTGAGGACAGCTGGGTGCCAGCCTCA 300

Db 318 CTGACCAGCTCACCAGCTACGTTTCTTCTCAGACTGAGGACAGCTGGTCCCGACCTCA 377  
QY 301 CCTCACAGGCTAGACCCCTCACAGTGGCGCCCATGTCCTCAGTGTGGTCTG 360  
Db 378 CCTCACAGGCTAGACCCCTCACAGTGGCGCCCATGTCCTCAGTGTGGTCTG 437  
QY 361 CTCAGCCGGGACCCCTCGGACC 382  
Db 438 CTCAGCCGGGDMCCCTGGACCC 459

## RESULT 2

US-09-513-999C-3603  
; Sequence 3603, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59, US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 3603  
; LENGTH: 303  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 68..301  
US-09-513-999C-3603

Query Match 48.0%; Score 236; DB 4; Length 303;  
Best Local Similarity 100.0%; Pred. No. 2.6e-58;  
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAGCTGCTCTGTAGACTCCAGCGCTTCCCAATGCCCCAG 60  
Db 68 ATGAGCACTACCGGAAGCTGCTCTGTAGACTCCAGCGCTTCCCAATGCCCCAG 127  
QY 61 CTACCTCCTGATACCTTGGATGCGGTCGAGATGCGAGCAAAATTCGCAACCTGTG 120  
Db 128 CTACCTCCTGATACCTTGGATGCGGTCGAGATGCGAGCAAAATTCGCAACCTGTG 187  
QY 121 GGTTGGCTTGGTGGTGGAGGCGGAGTGTCTCGCATGTAGTGTCTCAGTTCT 180  
Db 188 GGTTGGCTTGGTGGTGGAGGCGGAGTGTCTCGCATGTAGTGTCTCAGTTCT 247  
QY 181 GGCAGGCTGCAGGAAGCTGTACCTGGCTGAGTGTCAAGCGGGGTCCC 236  
Db 248 GGCAGGCTGCAGGAAGCTGTACCTGGCTGAGTGTCAAGCGGGGTCCC 303

## RESULT 3

US-09-620-312D-5  
; Sequence 5, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Felyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunging  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillingshast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP28  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 5  
; LENGTH: 1941  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (138)..(773)  
US-09-620-312D-5

Query Match 15.2%; Score 74.6; DB 4; Length 1941;  
Best Local Similarity 57.5%; Pred. No. 7.8e-12;  
Matches 134; Conservative 0; Mismatches 99; Indels 0; Gaps 0;  
QY 36 CCCAGCGCTTCCCAATGCCAGCTACCTCTGTATACCTTGAGATGCGGGTCCGAGA 95  
Db 239 CCCGGGCTCCGGCCCGTTTCGACAGACCTGGCGCCGCGGTGCACATGCGGGTCAAGGA 298  
QY 96 TGGCAGCAAAATTCGCAACCTGCTGGGTTGGCTCTGCTGCTGGAGGCGCGCAGTGC 155  
Db 299 AGGAGCAAGATCGGAACCTGTGAGCTTCGCCACCGCCAGCATGGCGCAGCCAGCCAC 358  
QY 156 TCGGCATGTAGTGTCTCAGGTTCTGCGAGGGCTGCAGGAAAGCTGTGACGTGCGCTGA 215  
Db 359 GCGGCCATCTCTTCAGCGGCTGCGGCGGGCCACACCAAAACCGTCAAGTGGCGGA 418  
QY 216 GATTCTCAAGCGGGGTCCCGAGGCTGCACCACTCACCAGTACGTATTC 268  
Db 419 GATCTCAAGCGCGCTTGGCGGCGCTGCACAGGTCAAGCGGCTGCGCTACC 471

## RESULT 4

US-09-976-594-92  
; Sequence 92, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 92  
; LENGTH: 2375  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: incyte ID No. 6673549 337221.11  
US-09-976-594-92

Query Match 15.2%; Score 74.6; DB 4; Length 2375;







```
RESULT 10
US-09-259-821A-1
; Sequence 1, Application US/09259821A
; Patent No. 6210926
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317
; CURRENT APPLICATION NUMBER: US/09/259,821A
; CURRENT FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-259-821A-1

Query Match      7.2%; Score 35.4; DB 3; Length 4257;
Best Local Similarity 52.3%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 326 CAGTGGCGCGCATGTGCTGAGTGGGTGCTCAGCGGGACCCCTGGACCCCA 385
    |||||
Db 2284 CTGCGGGGACCTGCGGTGGCGGCGGAGGCGCGGTGGCGGCGCC 2343

QY 386 ATGAGTGTGTTACCAACCCCGAGGAGCACCCTGGCTTCCATGCCAGCTCCA 445
    |||||
Db 2344 GTGAGCTGTGTCGCGGGGCGCTGGGCGCGCGTGGCGGACCCGCGCTGCCGAGC 2403

QY 446 GCTGTGCGCCCTCGTTCGCCGAAGAAGGCT 474
    |||||
Db 2404 TCCGCGCGCGCGCGCGCGGACCTGCT 2432

RESULT 11
US-08-843-659-1
; Sequence 1, Application US/08843659
; Patent No. 6218103
; GENERAL INFORMATION:
; APPLICANT: Leopardi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS
; INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,659
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARS8:519
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-843-659-1

Query Match      7.2%; Score 35.4; DB 3; Length 4257;
Best Local Similarity 52.3%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 326 CAGTGGCGCGCATGTGCTGAGTGGGTGCTCAGCGGGACCCCTGGACCCCA 385
    |||||
Db 2284 CTGCGGGGACCTGCGGTGGCGGCGGAGGCGCGGTGGCGGCGCC 2343

QY 386 ATGAGTGTGTTACCAACCCCGAGGAGCACCCTGGCTTCCATGCCAGCTCCA 445
    |||||
Db 2344 GTGAGCTGTGTCGCGGGGCGCTGGGCGCGCGTGGCGGACCCGCGCTGCCGAGC 2403

QY 446 GCTGTGCGCCCTCGTTCGCCGAAGAAGGCT 474
    |||||
Db 2404 TCCGCGCGCGCGCGCGGACCTGCT 2432

RESULT 12
US-09-825-288A-1
; Sequence 1, Application US/09825288A
; Patent No. 6723511
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317USC1
; CURRENT APPLICATION NUMBER: US/09/825,288A
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 09/259,821
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-1

Query Match      7.2%; Score 35.4; DB 4; Length 4257;
Best Local Similarity 52.3%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 326 CAGTGGCGCGCATGTGCTGAGTGGGTGCTCAGCGGGACCCCTGGACCCCA 385
    |||||
Db 2284 CTGCGGGGACCTGCGGTGGCGGCGGAGGCGCGGTGGCGGCGCC 2343

QY 386 ATGAGTGTGTTACCAACCCCGAGGAGCACCCTGGCTTCCATGCCAGCTCCA 445
    |||||
Db 2344 GTGAGCTGTGTCGCGGGGCGCTGGGCGCGCGTGGCGGACCCGCGCTGCCGAGC 2403

QY 446 GCTGTGCGCCCTCGTTCGCCGAAGAAGGCT 474
    |||||
Db 2404 TCCGCGCGCGCGCGCGGACCTGCT 2432

RESULT 13
US-08-458-568A-11/c
; Sequence 11, Application US/08458568A
; Patent No. 5821339
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Priscilla A.
; APPLICANT: Yeh, Lily
; TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
```



Db 1089 CAACGTCTACGAGCGCC 1106

Search completed: January 31, 2005, 18:09:29  
Job time : 98 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:01 ; Search time 32.6 Seconds  
(without alignments)  
481.084 Million cell updates/sec

Title: US-10-057-813-14  
Perfect score: 861  
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....SMPSSCGPRRRRRADTRS 163

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	219.5	25.5	538	2 H86335	T20H2.2 protein -
2	105.5	12.3	369	2 F96788	protein T4012.22 l
3	86.5	10.0	350	2 C26984	probable regulator
4	85.5	9.9	906	2 T00039	hypothetical prote
5	84.5	9.8	368	2 T35150	probable glycosyl
6	84	9.8	397	2 T30168	hypothetical prote
7	84	9.8	634	2 T00388	hypothetical prote
8	82.5	9.6	556	2 D70940	probable PPE prote
9	81.5	9.5	1275	2 AD0332	probable membrane
10	81	9.4	367	2 C39590	TPA-induced protei
11	80.5	9.3	257	2 T10586	small nuclear ribo
12	80.5	9.3	495	1 P2MWB2	I2 protein - bovin
13	80	9.3	589	2 C70767	probable pknJ - My
14	79.5	9.2	1052	2 T00067	hypothetical prote
15	79	9.2	469	1 P2MWB	I2 protein - bovin
16	79	9.2	770	2 G70718	probable cation tr
17	78.5	9.1	359	2 T36443	probable penicilli
18	78.5	9.1	579	2 C75599	TerF-related prote
19	78.5	9.1	2605	2 T18552	saframycin Mxi syn
20	78	9.1	817	2 S51342	verprolin - yeast
21	78	9.1	827	2 H83217	probable transcrip
22	77.5	9.0	6420	2 T30283	polyketide synthas
23	76.5	8.9	467	2 T34874	hypothetical prote
24	76.5	8.9	474	2 T19543	hypothetical prote
25	76.5	8.9	1460	1 EDBEIF	immediate-early pr
26	76.5	8.9	2142	2 B35098	MHC class III hist
27	76	8.8	202	2 AF2088	hypothetical prote
28	76	8.8	434	1 WMBEF3	UL43 protein - hum
29	76	8.8	537	2 I53719	NP-kappa-B transcr

30	76	8.8	2774	2 A43359	microtubule-associ
31	76	8.8	3343	2 T42207	breast cancer ausc
32	75.5	8.8	231	2 D72532	hypothetical prote
33	75.5	8.8	405	2 T44488	arac-family transc
34	75.5	8.8	433	2 F75566	probable benzoste
35	75.5	8.8	466	2 AH3100	amidohydrolase l
36	75.5	8.8	466	2 B98186	probable hydrolase
37	75	8.7	271	2 S48426	RHR2 protein - yea
38	75	8.7	278	2 T31319	triose-phosphate i
39	75	8.7	307	2 B72677	hypothetical prote
40	75	8.7	560	2 D30930	flagellar basal bo
41	75	8.7	565	2 T08794	hypothetical prote
42	75	8.7	1446	1 A45344	immediate-early pr
43	75	8.7	1493	2 A38218	GAP-associated pro
44	75	8.7	7463	2 T36248	CDA peptide synth
45	74.5	8.7	336	2 T17408	rRNA (adenine-N6-)

ALIGNMENTS

RESULT 1

H86335

T20H2.2 protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: H86335

R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: H86335

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-538 <STO>

A;Cross-references: UNIPROT:Q9LNU5; GB:AE005172; NID:98778978; PIDN:AAF79893.1; GSPDB:G

C;Genetics:

A;Map position: 1

Query Match 25.5%; Score 219.5; DB 2; Length 538;  
Best Local Similarity 30.4%; Pred. No. 4.6e-12;  
Matches 55; Conservative 31; Mismatches 64; Indels 31; Gaps 6;

QY	1	MEHYRKAGSVELP-APSPMPQLPPDLEMRVRDGSKIRNLGLALGRLEGSGSARHVVFSG	59
Db	224	MKIYQR---VEPKADTPIAE-----NEIRITSMGRARNYITVAMALLQNKNEVIFKA	275
QY	60	SGRAAGKAVSCAIEVRRVPGHLTKLRFLOTEDSWVPASPDITGLDPLTVRRHVPVWV	119
Db	276	MGRANKSVTVIVELIKRIPGLHQITSGTSDITDWTPT--EGLQTIETTRHVSMTI	333
QY	120	LISRDLDPNECYQPP-----GAPPLGSMPPSSSCGPRRRRRARDT	161
Db	334	TLSKEQLNTSSVGYQCPIPIEMVKPLAIEDYEQDGSPPRGRRGR--GGRGRGRGG	391
QY	162	R 162	
Db	392	R 392	

RESULT 2

F96788

protein T4012.22 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: F96788

R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:111130712  
A;Accession: F96788  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-369 <STO>  
A;Cross-references: UNIPROT:Q9LQRT; GB:AE005173; NID:g8778814; PIDN:AAF79819.1; GSPDB:GN C;Genetics:  
A;Gene: T4012.22.  
A;Map position: 1

Query Match 12.3%; Score 105.5; DB 2; Length 369;  
Best Local Similarity 22.5%; Pred. No. 0.055;  
Matches 34; Conservative 17; Mismatches 47; Indels 53; Gaps 4;  
QY 1 MEHYRKAGSVLPAPSPMPQLPPTLEMRVRDGSKIRNLGLALGRLE-----48  
Db 1 MDKYQR-----VVKPKADTPIDANEIRITSQGRARNYITAMTLLQVNESEMCOFTI 53  
QY 49 -----GGARHVFGSGRAAGKAVCAEIVK-----R 76  
Db 54 HCLVFAVACWEIFRDKGSTWFKAMGRAINTVTIVELIKVFFWFALSNRSRVQMGR 113  
QY 77 RVPGLHQLTKLRFLOQTESWVPASPDGLDP 107  
Db 114 RIPDLHQNVSIGSTDITDITWEPT--EGLLP 142

RESULT 3  
C26984  
probable regulatory protein 5 - Streptomyces griseus  
C;Species: Streptomyces griseus  
C;Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004  
C;Accession: C26984  
R;Distler, J.; Ebert, A.; Mansouri, K.; Pissowetzki, K.; Stockmann, M.; Piepersberg, W. Nucleic Acids Res. 15, 8041-8056, 1987  
A;Title: Gene cluster for streptomycin biosynthesis in Streptomyces griseus: nucleotide A;Reference number: A93676; MUID:88040426; PMID:3118332  
A;Accession: C26984  
A;Molecule type: DNA  
A;Residues: 1-350 <Dis>  
A;Cross-references: UNIPROT:P08076; GB:Y00459; GB:S55493; NID:g1621271; PIDN:CAA68515.1; C;Genetics:  
A;Gene: strR  
C;Keywords: transcription regulation

Query Match 10.0%; Score 86.5; DB 2; Length 350;  
Best Local Similarity 29.1%; Pred. No. 2.7;  
Matches 46; Conservative 11; Mismatches 52; Indels 49; Gaps 9;  
QY 1 MEHYR--KAGSVLPAPSPMPQLPPTLEMRVRD-----SKIRNLGLALGRLEGGSA 52  
Db 54 VEHIRTAAAGAEIPAIVMP-----TTKRVIDGMHRLRATKRGATEIAVRYFEGGEE 107  
QY 53 RHVFGSGRAAGKAVCAEIVKRVV-GL-----HQLTKLRFLOQTESWVPASPD--T 103  
Db 108 EAFIFA-----VKSNTHGLPLSLDRKAAATRVLETHPSWSDRAGLAT 152  
QY 104 GLDPLTV-----RRHVPVAVVLLSRD-----PLDPNE 130  
Db 153 GLSAKTGTLRSCSTAGVPSQSNVRIGRGRAPLDPE 190

RESULT 4

T00039

hypothetical protein KIAA0290 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 09-Jul-2004  
C;Accession: T00039  
R;Ohara, O.; Nagase, T.; Ishikawa, K.; Nakajima, D.; Ohira, M.; Seki, N.; Nomura, N. submitted to the EMBL Data Library, August 1997  
A;Description: Prediction of the coding sequences of unidentified human genes.  
A;Reference number: Z14073  
A;Accession: T00039  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-906 <OHA>  
A;Cross-references: UNIPROT:O14526; EMBL:AB006628; NID:d1170681; PIDN:BAA22959.1; PID:d1 A;Experimental source: brain  
C;Genetics:  
A;Note: KIAA0290

Query Match 9.9%; Score 85.5; DB 2; Length 906;  
Best Local Similarity 26.5%; Pred. No. 9.2;  
Matches 49; Conservative 17; Mismatches 60; Indels 59; Gaps 11;  
QY 13 PAPSPMPQLPDP--TLEMRVRDGSKIRNLGLALGRLEGG--ARHVFSGS-----BAA 64  
Db 380 PAPAPAPACSPAAQAQLRATAGS-----LILPPGGTGMKRHSSRDAGKPPRPSA 432  
QY 65 GKAVCAEIVKRVVLPGLHQLTK-----LRFLOQTESWVPASP-- 101  
Db 433 PTSSCAE-----RLQSEQVSKNLPGLPGLSFAFDHEDFTGSSSLGFTSGPPFSSSPEN 488  
QY 102 --DTGLDPLTVRRHV-----PAVWVLLSRDPLDNECYQPPGA-----PPGLGSMF--S 147  
Db 489 VEDSLDPS-----HAAPGSPDSWVVRPGTQSPSPSCRAPPEARIGIRAPLPDPSQPLA 545  
QY 148 SSCGP 152  
Db 546 SSPGP 550

RESULT 5

T35150  
probable glycosyl transferase - Streptomyces coelicolor (fragment)  
C;Species: Streptomyces coelicolor  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
C;Accession: T35150  
R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, July 1998  
A;Reference number: Z21570  
A;Accession: T35150  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-368 <SEE>  
A;Cross-references: UNIPROT:Q8CJM1; EMBL:AL031107; PIDN:CAA19930.1; GSPDB:GN00070; SCORE C;Genetics:  
A;Gene: SCOREDB:SCSA7.01

Query Match 9.8%; Score 84.5; DB 2; Length 368;  
Best Local Similarity 28.0%; Pred. No. 4.2;  
Matches 46; Conservative 11; Mismatches 54; Indels 53; Gaps 11;  
QY 12 LPAPSPMPQLPDPPTLEMRVRDGSKIRNLGLALGRLEGG--ARHVFSGSRAAGKAVSCA 71  
Db 47 LPVGDWPR--PD-----DAARAEARVRLRPDGA--V--VLDG-----LVACG 86  
QY 72 --EIVKRVVLPGLHQLTKLRFLOQTESWVPASPDGLDPL-----TVRRHVPV-- 117  
Db 87 VPEVV---VPEARLRMAVLVH-----LPLGDETLGDPVAVAGLDARETIVRAVPAVA 138  
QY 118 ---WV---LLSRDPLDNECYQPPGA-----PPGLGSMFSSSC 150  
Db 139 TSDWAVRRLVSHHGLDPGRVHVAAPGADIAPADPTDGVSKLVC 182



T30168  
hypothetical protein F26B1.2 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T30168  
R;Latreille, P.; Wameley, P.; Kramer, J.  
submitted to the EMBL Data Library, November 1996  
A;description: The sequence of *C. elegans* cosmid F26B1.  
A;Reference number: Z20748  
A;Accession: T30168  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-397 <LAT>  
A;Cross-references: UNIPROT:P91277; EMBL:U80444; PIDN:AAB37791.1; GSPDB:GN00019; CESP:F26B1  
A;Experimental source: strain Bristol N2; clone F26B1  
C/Genetics:  
A;Gene: CESP:F26B1.2  
A;Map position: 1  
A;Introns: 16/3; 40/2; 78/3; 93/2; 117/3; 320/3; 382/2  
Query Match 9.8%; Score 84; DB 2; Length 397;  
Best Local Similarity 26.5%; Pred. No. 5.1;  
Matches 35; Conservative 19; Mismatches 52; Indels 16; Gaps 4;  
QY 31 RDGSKIRNLGLALGRLEGGARSHVVFSGRAAGKAVSCAEIVKRRVPGLHQ--ITKLR 88  
DB 142 RNGSKIKEL-----REKCSARKIFTGCAPGSTDRVLITSGEQKNVLGIIEVMKELK 194  
QY 89 FLOTEDSWVPASPDGTGLDPLTVRRH-----VPAVWVLLSRDPLDPNECGYQPPGAPGL 142  
DB 195 EIPKGSATPYLPAFNYPDSNISDYGFPGNMPAGGPNRNGPA-PQRGGQPPGGRPSY 253  
QY 143 GSNPSSSCGPRS 154  
DB 254 GGAITQGGQRRS 265  
RESULT 7  
T00388  
hypothetical protein KIAA0616 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004  
C/Accession: T00388  
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, M.  
DNA Res. 5, 169-176, 1998  
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete  
A;Reference number: Z14142; MUID:98403880; PMID:9734811  
A;Accession: T00388  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-634 <ISH>  
A;Cross-references: UNIPROT:O75114; EMBL:AB014516; NID:G3327045; PIDN:BA31591.1; PID:G3327045  
A;Experimental source: brain  
C/Genetics:  
A;Gene: KIAA0616  
Query Match 9.8%; Score 84; DB 2; Length 634;  
Best Local Similarity 22.0%; Pred. No. 8.5;  
Matches 37; Conservative 23; Mismatches 76; Indels 32; Gaps 5;  
QY 9 SVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLGLALGRLEGGARSHVVFSGRAAGKAV 68  
DB 251 NIHPSPPLPLDPEEFTFFALSSSSSTGNL-----AANLTHIGICGACGGHSTPG 301  
QY 69 SCAEIVKRRVPGLHQLTKLRFLOTEDSWVPASPD---GLDPLTVRRHVPVWVLLS--- 122  
DB 302 SSPQHRPAGVSPLSLSTEARQQASPTLSPLSPITQAVAMDALSLEQLPYAFTQAGSQ 361  
QY 123 -----RDLDPNECGYQPP-----GAPPGLSGMPSSSC--GPR 153  
DB 362 QPPFPQPPPPPPPPASQPPPPPPPPQAPVRLPFGGPLLPASLTRGPO 409

```

A:Molecule type: DNA
A:Residues: 1-495 <GRO>
A:Cross-references: UNIPROT:P06457; GB:X01768; GB:M24326; NID:g60859
C:Superfamily: papillomavirus L2 protein
C:Keywords: glycoprotein; late protein
F:28,459,478/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          9.3%; Score 80.5; DB 1; Length 495;
Best Local Similarity 23.8%; Pred. NO. 13;
Matches 38; Conservative 21; Mismatches 62; Indels 39; Gaps

QY      4 YRKAGSVLPAPSPNPQLPPTLEMRVDSGKIRNLLGLAL-----GRL-EGG
      ||      ||      ||      ||      ||      ||      ||      ||      ||
Db      44 YRTCKOAGTCPPDPVPKVEGTI-----ADKILKGLGLAIYLGGLGIGTWTSTCRVAAGG

QY      51 SARHVVVFGSG-----RAAGKAVSCAEIVKRRVPG-----LHQTLKRLFLQTESDWSVPAS
      ||      ||      ||      ||      ||      ||      ||      ||      ||
Db      98 SFRYVPLRTSGSTTSLASVGSFAGATGTRSGITGIPDLTLETICALPGAYEDTVLPEA

QY      101 P-----DTGLDLPITVRRHVPVAVVLLSRDPLDENE 130
      ||      ||      ||      ||      ||      ||      ||      ||      ||
Db      158 PAIVTPDPAVDTGIDGLSIGDSSTETLTILLPEGED 197
      ||      ||      ||      ||      ||      ||      ||      ||      ||

```

RESULT 13  
C70767  
probable pknJ - Mycobacterium tuberculosis (strain H37Rv)  
C\_Species: Mycobacterium tuberculosis  
C\_Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C\_Accession: C70767

Probable pknJ - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: C70767  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; H  
R:Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; H  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: C70767  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-589 <COL>  
A:Cross-references: UNIPROT:Q10697; GB:Z73966; GB:AL123456; NID:g3261577; PIDN  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: pknJ  
C:Superfamily: Mycobacterium tuberculosis probable serine/threonine-specific p  
F13-276/Domain: protein kinase homology <KIN>

Gene: *glaK*  
CDS: Surfactant; Mycobacterium tuberculosis probable serine/threonine-specific protein kinase  
FJ3-276/Domain: protein kinase homology <KIN>

	Query Match	Best Local Similarity	Score 80;	DB 2;	Length 589;
	Matches	Conservative	25.6%; Pred. No. 18;	Mismatches	Indels
	44;	50;	Gaps	5;	
QY	23	PDTLEMRVRDGSKIRNLLGLRLGLEGGARHVVFGSGRA--AGKAVSCAEIVKRVPV	80		
	:	:	:	:	:
	:	:	:	:	:
Db	185	PEVLGAQGDFGRADLYSLGCAFLRLTGEAPPAAGAGA AVVAGHLHQPPPTVSDRVPG	244		
	:	:	:	:	:
	:	:	:	:	:
QY	81	L-----HQLTKLRFLOTEDSVVPASPDLTGLDPLTVRRH	113		
	:	:	:	:	:
	:	:	:	:	:
Db	245	LSAAMDVAVTATAMAKDPNRFTSAGEFAHAALYYGGATDGVWPPSP-----	292		
	:	:	:	:	:
	:	:	:	:	:
QY	114	VPAVMVLISRDLPDNECYQPPGA-----PPLGLSMFPSSSCGPRSRRAR	159		
	:	:	:	:	:
	:	:	:	:	:
Db	293	APHV---ISQGVAPGSPMWQHHPVGVSVALATPFGHGWPGLPLPRRPRRYR	341		
	:	:	:	:	:
	:	:	:	:	:

RESULT 14  
T00067  
hypothetical protein KIAA0453 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 09-Jul-2004  
C:Accession: T00067  
R:Seiki, N.; Ohira, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.  
DNA Res. 4, 345-349, 1997

Search completed: January 26, 2005, 15:07:31  
Job time : 34.6 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:40 ; Search time 174.603 Seconds  
(without alignments)  
537.138 Million cell updates/sec

Title: US-10-057-813-14  
Perfect score: 861  
Sequence: 1 MEHYRKGAGVELPAPMPQ.....SMPSSCGPRRRRRDTRTS 163

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	861	100.0	163	Q8N5L8	Q8N518 homo sapien
2	815	94.7	163	Q99JH1	Q99JH1 mus musculus
3	814	94.5	163	Q9CYX4	Q9CYX4 mus musculus
4	421	48.9	224	O6DGS1	O6DGS1 brachydanio
5	350	40.7	199	Q91WE3	Q91WE3 mus musculus
6	343	39.8	199	Q9BUL9	Q9BUL9 homo sapien
7	339	39.4	199	Q9NXX8	Q9NXX8 homo sapien
8	263	30.5	215	Q7PTH5	Q7PTH5 anopheles g
9	235.5	27.4	206	Q9V9B6	Q9V9B6 drosophila
10	230.5	26.8	206	Q6NNC7	Q6NNC7 drosophila
11	230.5	26.8	206	AAR96155	AAR96155 drosophila
12	219.5	25.5	315	Q944A2	Q944A2 arabidopsis
13	219.5	25.5	538	Q9LNU5	Q9LNU5 arabidopsis
14	206	23.9	350	Q93VA8	Q93VA8 arabidopsis
15	201.5	23.4	732	Q7XJ13	Q7XJ13 oryza sativ
16	200	23.2	262	Q7XT90	Q7XT90 oryza sativ
17	197.5	22.9	272	Q8HTC6	Q8HTC6 arabidopsis
18	192.5	22.4	485	Q9M8Z6	Q9M8Z6 arabidopsis
19	167	19.4	241	Q7RW66	Q7RW66 plasmodium
20	165	19.2	248	Q8IAX8	Q8IAX8 plasmodium
21	117	13.6	211	Q8IDN4	Q8IDN4 plasmodium
22	108	12.5	202	Q7RPX4	Q7RPX4 plasmodium
23	105.5	12.3	369	Q9LQR7	Q9LQR7 arabidopsis
24	104	12.1	206	Q7YWL9	Q7YWL9 caenorhabdi
25	100.5	11.7	1080	Q7RVU1	Q7RVU1 neurospora
26	100	11.6	1091	Q9C246	Q9C246 neurospora
27	91.5	10.6	275	Q9AWP6	Q9AWP6 oryza sativ
28	91.5	10.6	341	Q8ISG7	Q8ISG7 stytonychia
29	91	10.6	691	Q7ZUB6	Q7ZUB6 brachydanio
30	90.5	10.5	443	Q9FBM9	Q9FBM9 streptomyc
31	89.5	10.4	218	Q9FVZ8	Q9FVZ8 oryza sativ

32	88.5	10.3	215	1	CNO MOUSE
33	88	10.2	479	2	Q6ZAY3
34	88	10.2	479	2	BAC99602
35	88	10.2	525	2	Q94JH5
36	87.5	10.2	510	2	Q8BLJ9
37	87.5	10.2	625	2	Q93H39
38	87.5	10.2	782	1	Q93H39
39	87	10.1	416	2	Q9KZE7
40	87	10.1	711	2	Q6KAL9
41	87	10.1	711	2	BAD21438
42	87	10.1	1098	2	Q811E7
43	86.5	10.0	350	1	STRR_STRGR
44	86.5	10.0	806	2	Q8VD37
45	86.5	10.0	885	2	Q8KY52

ALIGNMENTS

RESULT 1

Q8N5L8 PRELIMINARY; PRT; 163 AA.  
AC Q8N5L8;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Chromosome 9 open reading frame 23 protein.  
GN Name=C9orf23;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=223388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.,  
RA Krzywicki M.I., Skalka U., Smaluk D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
EMBL: BC032136; AAH32136.1; -  
DR Genew; HGNC:19909; C9orf23.  
DR InterPro; IPR011574; Alba\_DUF78.  
DR ProDom; PD010497; Alba\_DUF78; 1.  
SQ SEQUENCE 163 AA; 17631 MW; 442C8727191A0BCE CRC64;

Query Match 100.0%; Score 861; DB 2; Length 163;

Best Local Similarity 100.0%; Pred. No. 1.5e-66;

Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKGAGVELPAPMPQPPDTLEMRVRCGSKIRNLGLALGLEGSGARHVFSGS 60  
|||||

Db 1 MEHYRKAGSVLPAPSPMPQLPDDTLEMRVROGSKIRNLLGLALGRLEGSGARHVVFSGS 60  
 QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVWVL 120  
 Db 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVWVL 120

QY 121 LSRDPLDNECGYQPPGAPGLGSMPSGSCGPRRRARDTRS 163  
 Db 121 LSRDPLDNECGYQPPGAPGLGSMPSGSCGPRRRARDTRS 163

## RESULT 2

Q99JH1 ID Q99JH1 PRELIMINARY; PRT; 163 AA.  
 AC Q99JH1  
 DT 01-JUN-2001 (TREMELrel. 17, Created)  
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
 DT 01-OCT-2004 (TREMELrel. 28, Last annotation update)  
 DE Hypothetical protein (RIKEN cDNA 2810432D09).  
 GN Name=2810432D09Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RA Ievolella C. Zara I., Lanfranchi G.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N-3;  
 RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by  
 RC biopsy.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N-3;  
 RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by  
 RC biopsy.  
 RA Strausberg R.;

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ293897; CAC34588.1; -;  
 DR EMBL; BC047068; AAH47068.1; -;  
 DR MGD; MGI:1917211; 2810432D09Rik.  
 DR InterPro; IPR011574; Alba DUF78.  
 DR ProDom; PD010497; Alba\_DUF78; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 163 AA; 17675 MW; 87675201AF87F5B6 CRC64;

Query Match 94.7%; Score 815; DB 2; Length 163;  
 Best Local Similarity 94.5%; Pred. No. 1.4e-62;  
 Matches 154; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPQLPDDTLEMRVROGSKIRNLLGLALGRLEGSGARHVVFSGS 60  
 Db 1 MEHYRKAGSVLPAPSPMPQLPDDTLEMRVROGSKIRNLLGLALGRLEGSGARHVVFSGS 60  
 QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVWVL 120  
 Db 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVWVL 120  
 QY 121 LSRDPLDNECGYQPPGAPGLGSMPSGSCGPRRRARDTRS 163  
 Db 121 LSRDPLDNECGYQPPGAPGLGSMPSGSCGPRRRARDTRS 163

## RESULT 3

Q9CYX4 ID Q9CYX4 PRELIMINARY; PRT; 163 AA.  
 AC Q9CYX4  
 DT 01-JUN-2001 (TREMELrel. 17, Created)  
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMELrel. 24, Last annotation update)  
 DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length  
 DE enriched library, clone:2810432D09 product:hypothetical protein, full  
 DE insert sequence.  
 GN Name=2810432D09Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RL "High-efficiency full-length cDNA cloning.";  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 409:685-690 (2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630 (2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Kiteunai T., Tashiro H., Itoh M.,  
 RA Konno H., Akiyama J., Nishi K., Nishine T., Harada A.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Ikegami T., Kashiwagi K.,  
 RA Yamamoto R., Mateumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";

Genome Res. 10:1757-1771 (2000).

[6]

RL STRAIN=C57BL/6J; TISSUE=Whole body;  
 RN SEQUENCE FROM N.A.  
 RC Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Imotani K., Ishii Y., Itoh M., Iwano M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sogabe Y., Suzuki H., Shibata K., Shibata Y., Shingawa A., Shiraki T.,  
 RA Tejima Y., Terauchi H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Terauchi H., Toyama T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Yamanaka M., Hayashizaki Y.,  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK013222; BAB28723.1; -  
 DR MGI; 1917211; 2810432D09R1k.  
 DR InterPro; IPR011574; Alba\_DUF78.  
 DR ProDom; PD010497; Alba\_DUF78; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 163 AA; 17648 MW; DC89421F4F617429 CRC64;

Query Match 94.5%; Score 814; DB 2; Length 163;  
 Best Local Similarity 94.5%; Pred. No. 1.8e-62;  
 Matches 154; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MEHYRKAGSVLPAPSPMPQPPDTLEMRVDRGSKIRNLLGLALGRLEGSGSARHVFGS 60  
 Db 1 MEQYREAGSVLPASSMPQPPDTLEMRVDRGSKIRNLLGLALGRLEGSGSTRHVFGS 60

Qy 61 GRAAGKAVSCAIEVKRRVPGHQLTKLRFQTEDSWVPASPDGTGLDPLTVRRHVPVAVWL 120  
 Db 61 GRAAGKAVSCAIEVKRRVPGHQLTKLRFQTEDSWVPTSPDTGLDPLTVRRHVPVAVWL 120

Qy 121 LSRDPLDPNCGYPPGAPGLSGMPSSSCGPRRRARDTRS 163  
 Db 121 LSRDPLDPNCGYPPGAPGLSGIPSPGPRRRARDTRS 163

RESULT 4  
 Q6DGS1 ID Q6DGS1 PRELIMINARY; PRT; 224 AA.  
 AC Q6DGS1  
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DE 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein.  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Jones S.J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RL "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC076268; AAH76268.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 224 AA; 24369 MW; C2FD38C9D8B2F38A CRC64;

Query Match 48.9%; Score 421; DB 2; Length 224;  
 Best Local Similarity 47.5%; Pred. No. 2.2e-28;  
 Matches 86; Conservative 27; Mismatches 46; Indels 22; Gaps 3;

Qy 1 MEHYRKAGSVLPAPSPMPQPPDTLEMRVDRGSKIRNLLGLALGRLEGSGS----- 52  
 Db 1 MENYRKANTTIEQPCPCPPFDLPDPSDTPVRVKDGSKIRNLMRFLSRMEESTAAADHEGS 60

Qy 53 -----RHVVFSGSGRAAGKAVSCAIEVKRRVPGHQLTKLRFQTEDSWVPASP 101  
 Db 61 EVSVGTGDNLCROIQVFTGCGQSVAKAITCVMKRIHGHQLTKLRYRTLDQVWEPLEP 120

Qy 102 DTGLDPLTVRRHVPVAVWLSDRDPNCGYPPGAPGL--GSMFSSSCGPR--SRRA 158  
 Db 121 GAGLDSLTYSRVNPSVWLLSRDLSLXNPGYQAPGSDALWQALKEAATPRHGRRK 180

Qy 159 R 159  
 Db 181 R 181

RESULT 5  
 Q91WE3 ID Q91WE3 PRELIMINARY; PRT; 199 AA.  
 AC Q91WE3  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DE 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Ribonuclease P 25KDa subunit (Mus musculus ES cells cDNA, RIKEN full-  
 DE length enriched library, clone:C33003C13 product:hypothetical  
 DE protein, full insert sequence).  
 GN Name=Rpp25;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Jones S.J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RL "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).





```

DR EMBL: AY034074; AAK54443.1; --
DR Interact; Q9BUL9; --
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 199 AA; 20632 MW; 9A4494F0297B2A81 CRC64;

Query Match      39.8%; Score 343; DB 2; Length 199;
Best Local Similarity 42.4%; Pred. No. 1.1e-21;
Matches 73; Conservative 26; Mismatches 55; Indels 18; Gaps 3;

Qy 1 MEHYRKAGSVLPAP-----PSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG 50
Dy 1 MENFRKVSSEAPAGCGAGGPGSPADLAPGAVHVRVKEGSKIRNLMFATASWAQP 60
Qy 51 SARHVFSGSRAAGKAVSCAIEVKRRVPGHLQTLKFLQTEDSW--VPASPDGTGLDP- 107
Dy 1 ATRAIVFSGCGRATTKVTCAEILKRLAGLHQVTLRYRSVREVWQSLPPGPTQGTG 120
Qy 108 -----LTVRRHVPVAVVLLSRDPLDNECGYPPGAPGLGSMFSSSCGPRS 154
Dy 121 EPAASLSVLKNVPLGLAILLSKDALDPRQPGYQPPNPHFGPSPPAAPASKRS 172

RESULT 7
Q9NX88 PRELIMINARY; PRT; 199 AA.
ID Q9NX88 AC Q9NX88
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ20374.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
RA Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
RA Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK000381; BA911128.1; --
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 199 AA; 20660 MW; 855494F036B2291 CRC64;

Query Match      39.4%; Score 339; DB 2; Length 199;
Best Local Similarity 41.9%; Pred. No. 2.3e-21;
Matches 72; Conservative 26; Mismatches 56; Indels 18; Gaps 3;

Qy 1 MEHYRKAGSVLPAP-----PSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG 50
Dy 1 MENFRKVSSEAPAGCGAGGPGSPADLAPGAVHVRVKEGSKIRNLMFATASWAQP 60
Qy 51 SARHVFSGSRAAGKAVSCAIEVKRRVPGHLQTLKFLQTEDSW--VPASPDGTGLDP- 107
Dy 61 ATRAIVFSGCGRATTKVTCAEILKRLAGLHQVTLRYRSVREVWQSLPPGPTQGTG 120
Qy 108 -----LTVRRHVPVAVVLLSRDPLDNECGYPPGAPGLGSMFSSSCGPRS 154
Dy 121 EPAASLSVLKNVPLGLAILLSKDALDPRQPGYQPPNPHFGPSPPAAPASKRS 172

RESULT 8
Q7PHT5 PRELIMINARY; PRT; 215 AA.
ID Q7PHT5 AC Q7PHT5
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000016002 (Fragment).
GN Name=ENSANGG00000013513;
OS Anopheles gambiae str. PEST.

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008807; EAA03913.2; --
FT NON TER 1 1
FT 215
SQ SEQUENCE 215 AA; 24004 MW; AA72B25FAA8ABFBE CRC64;

Query Match      30.5%; Score 263; DB 2; Length 215;
Best Local Similarity 39.6%; Pred. No. 9.4e-15;
Matches 65; Conservative 28; Mismatches 61; Indels 10; Gaps 5;

Qy 1 MEHYRKAGSVLPAP-----PMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG SARHVP 57
Dy 6 MMHYKGNVVEELSQQIFIEVLPAFLMHWKGSSEVKNLVDYAKKALEEGTHRSVVM 65
Qy 58 SSGSRAAGKAVSCAIEVKRRVPGHLQTLK--RFLQTEDSWVPASPDGTGLDPLTVRRHP 115
Dy 66 SSGDGGVGKTIACAEIMKRHFE-LHQVTRICYRKCVCVEEFWDP--QDEGLEQIVAKRNIP 122
Qy 116 AVWVLLSRDPLDNECGYPPGAPGL--GSMFSSSCGPRSRRR 157
Dy 123 CVHILMSLDBIDPSVAGYQHSKTQGGFWTGAAGLSSDGDPRKRTK 166

RESULT 9
Q9V9B6 PRELIMINARY; PRT; 206 AA.
ID Q9V9B6 AC Q9V9B6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE C99422-PA (C99422-pc).
GN ORFNames=C99422;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophiliidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., M.G.,  
RA Palazotto M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reineck K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Sinden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou G., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celisner S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
RT *melanogaster* euchromatic genome sequence.";  
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celisner S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genomic perspective.";  
RT Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celisner S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review.";  
RT Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RN SEQUENCE FROM N.A.  
RX FLYBASE;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RN SEQUENCE FROM N.A.  
RX FLYBASE;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003790; AA070814.1; -;  
DR IntAct; Q9V9B6; -;  
DR FlyBase; FBgn0033092; CG9422.  
SQ SEQUENCE 206 AA; 23146 MW; 9FCDF4C3C9BC5D9 CRC64;

Query Match 27.4%; Score 235.5; DB 2; Length 206;  
Best Local Similarity 33.0%; Pred. No. 2.1e-12;  
Matches 63; Conservative 27; Mismatches 70; Indels 31; Gaps 5;  
QY 1 MEHYRAGSVE-----LPAPSPMPQLPPDTPLEMRVDRGSKIRNLGLALGRLEGSSAR 53  
DB 1 MMHYRKAENVEKELSKDLFPEDCPKSKQDPLMHVKGKTVSNVIEFAQALNKGHR 60  
QY 54 HVPFSGSRAAGKAVSCAIEIVKRVPGHQLTKLRFLOTEDSVVPASPTDGLDPLTVRRH 113  
DB 61 CVVWSGGGGVKTICCAEVLKRSHPLVQVTRMAYTSVEHWKPM--EGLEEIIIVTRQ 117

QY 114 VPAVWLLSRDPLDNECGYQPPGAP-----PGLGSM-----PSSSCGP 152  
DB 118 IPTLHLSLDLPTIDGLQKPTSTDFWGGGGAQQQPHRSPRHHQPHKPGAGRG 177  
QY 153 RSRRRARDTRS 163  
DB 178 RPNKTRPGRN 188  
RESULT 10  
Q6NNC7 PRELIMINARY; PRT; 206 AA.  
AC Q6NNC7;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE RE65722P.  
GN Name=CG9422;  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; *Drosophila*.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,  
RA Park S., Wan K., Yu C., Rubin G.M., Celisner S.;  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BT011363; AAR96155.1; -;  
SQ SEQUENCE 206 AA; 23162 MW; 8B9FD01C279BC5D9 CRC64;  
Query Match 26.8%; Score 230.5; DB 2; Length 206;  
Best Local Similarity 32.5%; Pred. No. 5.7e-12;  
Matches 62; Conservative 27; Mismatches 71; Indels 31; Gaps 5;  
QY 1 MEHYRAGSVE-----LPAPSPMPQLPPDTPLEMRVDRGSKIRNLGLALGRLEGSSAR 53  
DB 1 MMHYRKAENVEKELSKDLFPEDCPKSKQDPLMHVKGKTVSNVIEFAQALNKGHR 60  
QY 54 HVPFSGSRAAGKAVSCAIEIVKRVPGHQLTKLRFLOTEDSVVPASPTDGLDPLTVRRH 113  
DB 61 CVVWSGGGGVKTICCAEVLKRSHPLVQVTRMAYTSVEHWKPM--EGLEEIIIVTRQ 117  
QY 114 VPAVWLLSRDPLDNECGYQPPGAP-----PGLGSM-----PSSSCGP 152  
DB 118 IPTLHLSLDLPTIDGLQKPTSTDFWGGGGAQQQPHRSPRHHQPHKPGAGRG 177  
QY 153 RSRRRARDTRS 163  
DB 178 RPNKTRPGRN 188  
RESULT 11  
AAR96155 PRELIMINARY; PRT; 206 AA.  
AC AAR96155;  
DT 02-MAR-2004 (TReMBLrel. 27, Created)  
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)  
DE RE65722P.  
GN CG9422.  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; *Drosophila*.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,  
RA Park S., Wan K., Yu C., Rubin G.M., Celisner S.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RN SEQUENCE FROM N.A.  
RX FLYBASE;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RN SEQUENCE FROM N.A.  
RX FLYBASE;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003790; AA070814.1; -;  
DR IntAct; Q9V9B6; -;  
DR FlyBase; FBgn0033092; CG9422.  
SQ SEQUENCE 206 AA; 23146 MW; 9FCDF4C3C9BC5D9 CRC64;



RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,  
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,  
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,  
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,  
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,  
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA EMBL; AF428441; AAL16210.1; -  
 DR EMBL; AF054208; AAL06869.1; -  
 DR EMBL; AF412102; AAL06555.1; -  
 DR EMBL; AY124847; AAM70556.1; -  
 DR HSP; P74761; IH0X.  
 DR InterPro; IPR011574; Alba DUF78.  
 DR ProDom; PD010497; Alba DUF78; 1.  
 SQ SEQUENCE 350 AA; 37363 MW; DB474865DF4E162E CRC64;  
 Query Match 23.9%; Score 206; DB 2; Length 350;  
 Best Local Similarity 29.9%; Pred. No. 1.3e-09;  
 Matches 53; Conservative 23; Mismatches 65; Indels 36; Gaps 4;  
 QY 1 MEHYKAGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLGLALGRLEGGSGARHVVFSGS 60  
 DB 1 MDYQYR-----VVKPADTPIDANEIRITSGQARNYITYATLLQDKGSTEVEVFKAM 53  
 QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLQTEDSWVPASPDGLDPLTVRRHVPVWVL 120  
 DB 54 GRAINKTVIVELIKRIIPDLHQNTSGTDTIDTWEP--EEGLPLETTRHVSMTIT 111  
 QY 121 LSRDPLDNECYQPP-----GAPPLGSPSSSCGPRRRAR 159  
 DB 112 LSKIELNTSSVGYQCPPIELVPMGDIDYEGREGSPGGRG-----RGRGR 159  
 RESULT 15  
 Q7XJ13 PRELIMINARY; PRT; 732 AA.  
 AC Q7XJ13;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Chloroplast nucleoid DNA-binding protein-like protein.  
 GN Name=F0478E02.13;  
 OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_taxid=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Oliveira A.C., Mattos L.T., Carvalho F.F., Shimano A., Zimmer P.D.,  
 RA Malone G., Dellagostin O.;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB109206; BAC79194.1; -  
 DR Gramene; Q7XJ13; -  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004194; F:pepsin A activity; IEA.  
 DR GO; GO:0006198; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR011574; Alba DUF78.  
 DR InterPro; IPR001461; Peptidase\_A1.  
 DR InterPro; IPR001969; Pept\_Asp\_AS.  
 DR Pfam; PF00026; Asp; 1.  
 DR ProDom; PD010497; ASP\_PROTEASE; UNKNOWN\_2.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 79; Indels 25; Gaps 3;  
 KW DNA-binding.  
 SQ SEQUENCE 732 AA; 78353 MW; 7EA86DD5386374AC CRC64;  
 Query Match 23.4%; Score 201.5; DB 2; Length 732;  
 Best Local Similarity 30.9%; Pred. No. 7.4e-09;  
 Matches 55; Conservative 19; Mismatches 79; Indels 25; Gaps 3;  
 QY 1 MEHYKAGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLGLALGRLEGGSGARHVVFSGS 60  
 DB 491 MDYQYR-----VEKPREAPIKENEIRITTOGRMRNYITYATLLQDKGSTEVEVFKAM 543  
 QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLQTEDSWVPASPDGLDPLTVRRHVPVWVL 120  
 DB 544 GRAINKTVIAELIKRIIVGLHQNTTGTDTIDTWEP--EEGLPLETTRHVSMTIT 601  
 QY 121 LSRDPLDNECYQPP-----GAPPLGSPSSSCGPRRRARDTR 162  
 DB 602 LSKKELDTSSIGYQSPFLPADKVKPLVEYNEEDAPSPAGRGGRGGGGRGRGRGTR 659  
 Search completed: January 26, 2005, 15:06:25  
 Job time : 177.603 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:47 ; Search time 41.4407 Seconds  
(without alignments)  
260.851 Million cell updates/sec

Title: US-10-057-813-14  
Perfect score: 861  
Sequence: 1 MEHYRKAGSVLPAPMPQ.....SMPSSCGPRRRARDTRS 163

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfilese1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	642	74.6	128	US-09-621-976-4077	Sequence 4077, Ap
2	395	45.9	78	US-09-513-999C-7680	Sequence 7680, Ap
3	151.5	17.6	118	US-09-270-767-39818	Sequence 39818, A
4	151.5	17.6	118	US-09-270-767-55035	Sequence 55035, A
5	91.5	10.6	950	US-09-252-991A-25927	Sequence 25927, A
6	91	10.6	345	US-09-252-991A-18806	Sequence 18806, A
7	91	10.6	921	US-09-252-991A-17504	Sequence 17504, A
8	90.5	10.5	330	US-09-252-991A-19363	Sequence 19363, A
9	90	10.5	429	US-09-252-991A-20363	Sequence 20363, A
10	89.5	10.4	432	US-09-252-991A-23622	Sequence 23622, A
11	87	10.1	374	US-09-252-991A-28830	Sequence 28830, A
12	85.5	9.9	446	US-09-252-991A-27110	Sequence 27110, A
13	85.5	9.9	477	US-09-252-991A-19831	Sequence 19831, A
14	85.5	9.9	497	US-09-252-991A-23620	Sequence 23620, A
15	85	9.9	441	US-09-252-991A-28965	Sequence 28965, A
16	84.5	9.8	282	US-09-252-991A-19978	Sequence 19978, A
17	84	9.8	546	US-09-252-991A-25851	Sequence 25851, A
18	83.5	9.7	1482	US-09-410-551B-21	Sequence 21, Appl
19	83.5	9.7	1482	US-09-940-316B-21	Sequence 21, Appl
20	83.5	9.7	1488	US-09-410-551B-17	Sequence 17, Appl
21	83.5	9.7	1488	US-09-940-316B-17	Sequence 17, Appl
22	83.5	9.7	1509	US-09-410-551B-23	Sequence 23, Appl
23	83.5	9.7	1509	US-09-940-316B-23	Sequence 23, Appl
24	83.5	9.7	1517	US-09-410-551B-19	Sequence 19, Appl
25	83.5	9.7	1517	US-09-940-316B-19	Sequence 19, Appl
26	83.5	9.7	6396	US-09-410-551B-72	Sequence 72, Appl
27	83.5	9.7	6396	US-09-940-316B-72	Sequence 72, Appl

28	83	9.6	240	4	US-09-252-991A-31100	Sequence 31100, A
29	83	9.6	594	4	US-09-252-991A-32578	Sequence 32578, A
30	82.5	9.6	472	4	US-09-252-991A-16723	Sequence 16723, A
31	82	9.5	404	4	US-09-252-991A-29985	Sequence 29985, A
32	81.5	9.5	469	4	US-09-252-991A-32604	Sequence 32604, A
33	81.5	9.5	491	4	US-09-252-991A-31758	Sequence 31758, A
34	81	9.4	302	4	US-09-252-991A-25724	Sequence 25724, A
35	81	9.4	482	4	US-09-252-991A-29113	Sequence 29113, A
36	81	9.4	516	4	US-09-252-991A-19788	Sequence 19788, A
37	80.5	9.3	266	4	US-09-252-991A-30538	Sequence 30538, A
38	80.5	9.3	964	3	US-08-484-791-2	Sequence 2, Appli
39	80	9.3	313	4	US-09-252-991A-23631	Sequence 23631, A
40	80	9.3	750	4	US-09-252-991A-23762	Sequence 23762, A
41	80	9.3	860	4	US-09-252-991A-28607	Sequence 28607, A
42	79.5	9.2	251	4	US-09-252-991A-31800	Sequence 31800, A
43	79.5	9.2	377	4	US-09-252-991A-24675	Sequence 24675, A
44	79.5	9.2	400	4	US-09-252-991A-31900	Sequence 31900, A
45	79.5	9.2	416	4	US-09-252-991A-30923	Sequence 30923, A

ALIGNMENTS

RESULT 1  
US-09-621-976-4077  
; Sequence 4077, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET 054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 4077  
; LENGTH: 128  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -52..-1  
; NAME/KEY: UNSURE  
; LOCATION: 72  
; OTHER INFORMATION: Xaa = Asp.Glu.Gly.Val  
US-09-621-976-4077

Query Match 74.6%; Score 642; DB 4; Length 128;  
Best Local Similarity 99.2%; Pred. No. 4.5e-62;  
Matches 124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MEHYRKAGSVLPAPMPQPPDTLEMRVDRGSKIRNLLGLALGEGSARHVVFSGS 60  
Db 1 MEHYRKAGSVLPAPMPQPPDTLEMRVDRGSKIRNLLGLALGEGSARHVVFSGS 60  
QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSVWPASPDGLDPLTVRRHVPAAVWL 120  
Db 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSVWPASPDGLDPLTVRRHVPAAVWL 120  
QY 121 LSRDP 125  
Db 121 LSRXP 125

RESULT 2  
US-09-513-999C-7680  
; Sequence 7680, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.

```
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7680
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-513-999C-7680

Query Match      45.9%; Score 395; DB 4; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.6e-35;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPOLPPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 1 MEHYRKAGSVLPAPSPMPOLPPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60

QY 61 GRAAGKAVCAEIVKRRV 78
Db 61 GRAAGKAVCAEIVKRRV 78

RESULT 3
US-09-270-767-39818
; Sequence 39818, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39818
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39818

Query Match      17.6%; Score 151.5; DB 4; Length 118;
Best Local Similarity 41.8%; Pred. No. 7.6e-09;
Matches 33; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 11 ELPAAPSPMPOLPPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGSRAAGKAVSC 70
Db 1 DLPPEDCMPSKQDFLWMHVKGTKVSNVIEFAQEALNKGHRVCVWSSGGSGGVGKTISC 60

QY 71 AEIVKRRVPGHLQTLKRF 89
Db 61 AEVLKRSHP-LYQVTRMAY 78

RESULT 4
US-09-270-767-55035
; Sequence 55035, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55035
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-55035

Query Match      17.6%; Score 151.5; DB 4; Length 118;
Best Local Similarity 41.8%; Pred. No. 7.6e-09;
Matches 33; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 11 ELPAAPSPMPOLPPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGSRAAGKAVSC 70
Db 1 DLPPEDCMPSKQDFLWMHVKGTKVSNVIEFAQEALNKGHRVCVWSSGGSGGVGKTISC 60

QY 71 AEIVKRRVPGHLQTLKRF 89
Db 61 AEVLKRSHP-LYQVTRMAY 78

RESULT 5
US-09-252-991A-25927
; Sequence 25927, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25927
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-25927

Query Match      10.6%; Score 91.5; DB 4; Length 950;
Best Local Similarity 25.8%; Pred. No. 0.39;
Matches 40; Conservative 12; Mismatches 60; Indels 43; Gaps 6;

QY 15 PSPMPOLPPTLEMRVRDGSKIRNLLGLALG--RLEGGSARHVVFSGSRAAGKAVSCAE 72
Db 338 PRPAPQLL-----YTRLRTALGAPAGMERLRGG---HCLQFATGAALGTEFGLPA 384

QY 73 IVKRRV-----PGLHQLTKLRFLOQTEDSVVPASPDGLDPLTVRRHVPVAVVLLSR- 123
Db 385 QVRRSDPAAARQRCGV-----ATGDPGRPATQRTAPGLTGRPOH 427

QY 124 --DPLDPNCGYQPFCAPEGLGSMFSSCGPSRR 156
Db 428 PGKPLRPGDAGHPQGRQPAERPELPLRPGKL 462

RESULT 6
US-09-252-991A-18806
; Sequence 18806, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```

; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18806  
; LENGTH: 345  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18806

Query Match 10.6%; Score 91; DB 4; Length 345;  
Best Local Similarity 25.7%; Pred. No. 0.12; Mismatches 12; Indels 80; Gaps 8;  
Matches 45; Conservative 12; Mismatches 12; Indels 80; Gaps 8;  
QY 4 YRKAGSVELPA-----PSMP-----QL--PPDTLEMRVD-----GSKIRNLLGL 42  
DB 228 HRGPGSQRLPAGRSRRRLFPVQPGAGQLLRPAELPRLRNGNGTEPAGTSGQNP GK 287  
QY 43 ALGRLEGGSARHVFGSGRAAGKAVSCAIEIVKRVVGLHQLTKLRFLQTEDSWVPASPD 102  
DB 288 A-PELSGQGRHLRDAAGRRSGLA-----PATPA 316  
QY 103 TGLDPLTVRRHVPAVMVLLSRDPLDNECGYQPPGAPGLGSMPSGCGPRSR 157  
DB 317 AGLPAATGR-----GRPGLGAVPLS-----PAARR 343

RESULT 7  
US-09-252-991A-17504  
; Sequence 17504, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17504  
; LENGTH: 921  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (886)  
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

US-09-252-991A-17504  
Query Match 10.6%; Score 91; DB 4; Length 921;  
Best Local Similarity 28.2%; Pred. No. 0.42;  
Matches 57; Conservative 11; Mismatches 70; Indels 64; Gaps 11;  
QY 11 ELPAPSPMPOL-----PPDTLEMRVDGSKIRNLLGLALG-RLEGGSARHVFGS 59  
DB 164 DLFPDRRGAALQGGRVPQVPHPHRRQVR-----RRLPLHAGLRQEGVHRH----- 213  
QY 60 SGRAAGKAVSCAIEIVK-----RVFGLHQLTKLRFLQTEDSWVPASPD TGL---DPLTVRR 112  
DB 214 PGRPRPLLPAPEDHRRRLARFGLHRAALLRLRLQRHGRPRDDGLPLHGLQPAL 273  
QY 113 HVPA-----VWVLLSRDPLD-----PNECGYQ-----PPGAPPG----- 141  
DB 274 HQPVDRHVLAALAHQPVDLAARLPLHQPGRQPRQHLPLDLPQVPDHPAGRFVARRQLHLH 333  
QY 142 -LGSM-----PSSGCGPRSRRA 158  
DB 334 HLGRLARHVAGDRAGPRSRRA 355

## RESULT 8

US-09-252-991A-19363  
; Sequence 19363, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19363  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19363

Query Match 10.5%; Score 90.5; DB 4; Length 330;  
Best Local Similarity 27.3%; Pred. No. 0.12;  
Matches 44; Conservative 12; Mismatches 72; Indels 33; Gaps 7;

QY 5 RKAGSVELPAPSPMPOLPPDTLEMRVDGSKIRNLLGLALGRLEGGSARHVFGSGRAA 64  
DB 87 RAAGRLRIPRAPDPDRVAADA-ERRHRGRPHSPGGAVALAALPSDVGGH---RPGRRAP 142  
QY 65 GRAVSCAIEIVK-----RVFGLH---QLTKLRFLQTEDSW---VPASPD TGLDPLTVRRHVPAV 117  
DB 143 GQAPKQRRARQPRRPRRHRHAGSLQRRSGIPEGAWRGPRPGPPAGRSPPAPR----- 197  
QY 118 WVLLSRDPLDNECGYQPPGAPGLGSMPSGCGPRSRRA 158  
DB 198 -----PAERRAALRSAGQPP-----GDRSRRA 221

## RESULT 9

US-09-252-991A-20363  
; Sequence 20363, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20363  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20363

Query Match 10.5%; Score 90; DB 4; Length 429;  
Best Local Similarity 31.3%; Pred. No. 0.2;  
Matches 51; Conservative 11; Mismatches 71; Indels 30; Gaps 8;

QY 12 LPA-----PSPMPOLPPDTLEMRVDGSKIR-----NLLGLALGRLEGGSARHVFGS---G 61  
DB 119 LPAGLGGVGLALPAVDLHARRADLPRLRLVPLQQAQAHLGQGVHRRFAGGDLPLPC 178  
QY 62 RAAGKAVSCAIEIVKRVVGLHQLTKLRFLQTEDSWVPASPD TGLDPLTVRRHVPAVMVLL 121  
DB 179 RGAGR-----LHRRHPGGEQRLRRRFAGL-DRAVPAVLRLPGPD-----RRLHPARLHLA 226

QY	122	SRPLDNECYQPPGAPGLGS	-----MPSSSCGP-RSRR	156
DB	227	DHEDRRPAGAHAPGAPAGVGAAGDRHGQOPVDAAGFCRDR	-----	269
RESULT 10				
US-09-252-991A-23622				
; Sequence 23622, Application US/09252991A				
; Patent No. 6551795				
; GENERAL INFORMATION:				
; APPLICANT: Marc J. Rubenfield et al.				
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS				
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS				
; FILE REFERENCE: 107196.136				
; CURRENT APPLICATION NUMBER: US/09/252,991A				
; CURRENT FILING DATE: 1999-02-18				
; PRIOR APPLICATION NUMBER: US 60/074,788				
; PRIOR FILING DATE: 1998-02-18				
; PRIOR APPLICATION NUMBER: US 60/094,190				
; PRIOR FILING DATE: 1998-07-27				
; NUMBER OF SEQ ID NOS: 33142				
; SEQ ID NO 23622				
; LENGTH: 432				
; TYPE: PRT				
; ORGANISM: Pseudomonas aeruginosa				
US-09-252-991A-23622				
Query Match 10.4%; Score 89.5; DB 4; Length 432;				
Best Local Similarity 26.9%; Pred. No. 0.23;				
Matches 45; Conservative 13; Mismatches 68; Indels 41; Gaps 8;				
QY	15	PSPMPQLPPTLEMRVR	-----DGSKIRNLLGLALGRLEGGSGARHVVFSGGAAKAV	68
DB	4	PRAPRLPDLQPPVTRTHRADGRTLAEPGGAAGR	-----PAAAHGRIAAVRRRGAAAT	60
QY	69	SCAEIVKR-RVPGHLQHLTKLFLQTESWVPASPDITGLDP	---LTVRRHVPVAVVLLSRD	124
DB	61	AARRHGRGRAPG	-----HPPMPVRRPGGQPPGPGADLRRTDPE	106
QY	125	PLDPN	-----ECGYQPPCAPPGGLGSPSSCGPRSRRRR	159
DB	107	PLOPSGARPAQSRRRPGSAGRCPGAFAGDGRRRPAGD	--PQGRRLR	151
RESULT 11				
US-09-252-991A-28830				
; Sequence 28830, Application US/09252991A				
; Patent No. 6551795				
; GENERAL INFORMATION:				
; APPLICANT: Marc J. Rubenfield et al.				
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS				
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS				
; FILE REFERENCE: 107196.136				
; CURRENT APPLICATION NUMBER: US/09/252,991A				
; CURRENT FILING DATE: 1999-02-18				
; PRIOR APPLICATION NUMBER: US 60/074,788				
; PRIOR FILING DATE: 1998-02-18				
; PRIOR APPLICATION NUMBER: US 60/094,190				
; PRIOR FILING DATE: 1998-07-27				
; NUMBER OF SEQ ID NOS: 33142				
; SEQ ID NO 28830				
; LENGTH: 374				
; TYPE: PRT				
; ORGANISM: Pseudomonas aeruginosa				
US-09-252-991A-28830				
Query Match 10.1%; Score 87; DB 4; Length 374;				
Best Local Similarity 24.6%; Pred. No. 0.35;				
Matches 41; Conservative 11; Mismatches 63; Indels 52; Gaps 5;				
QY	15	PSPMPQLPPTLEMRVR	DGSKIRNLLGLALGRLEGGSGARHVVFSGGAAKAVSCAEIV	74



```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19831

Query Match
Best Local Similarity 29.1%; Score 85.5; DB 4; Length 477;
Matches 39; Conservative 11; Mismatches 49; Indels 35; Gaps 7;

QY 43 ALGRLE-GGSAHVVFSGG-----RAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSW 96
DB 338 ATGRLHPGCGRRMARAVAGPREVPKAPAGAGGAGRRAPGLYQ----- 384
QY 97 VPASPTGLDP-----LTVRRHVPVWVLLSRDPLDNECGYQPPGAPGLGMPSS 148
DB 385 -PCAAQRGDPGRRPGAGRLGGORRRPG-YIAPGREPGDPR---TPAGALPGSGTAPG- 438
QY 149 SCGPRRRRRARDTR 162
DB 439 --GSHSRGRGRTPR 450

RESULT 14
US-09-252-991A-23620
; Sequence 23620, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23620
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23620

Query Match
Best Local Similarity 26.9%; Score 85.5; DB 4; Length 497;
Matches 49; Conservative 10; Mismatches 52; Indels 71; Gaps 11;

QY 19 POLPPDTLEMRV-----RDGSKIRNLLGLALGRLEG---GSAHVVFSGSGRAAGKAVSC 70
DB 244 PGQPPGFLPARQGRRRPGRAGSDLRAPFAATAGRSAGDRQGRRLL-----RAAGPG--- 293
QY 71 AEIVKRRVPGLHQLTKLRFLOTE-----DSW-----VP-----ASPDGLDP- 107
DB 294 ---LPQFTPGHARRORRRVAADPGSPVAPAPWGRGHADLPQOPQARQLAEPGTGPAPA 350
QY 108 -----LTVRRHVPVWVLLSRDPLDNECGYQPPGAPGLGMPSSCG---PRSR 156
DB 351 RROSRLRHLRLHRPA-----EGRGQHPRGA---GRAPAMDAGHLRPRRR 394
QY 157 RA 158
DB 395 RA 396

RESULT 15
US-09-252-991A-28965
; Sequence 28965, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
```

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28965
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (338),(413)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-28965

Query Match
Best Local Similarity 27.0%; Score 85; DB 4; Length 441;
Matches 47; Conservative 14; Mismatches 71; Indels 42; Gaps 8;

QY 5 RKAGSVELPAPSPMP-----QLPPDPTLEMRVDRDGSKIRNLLGLALGRLEG 49
DB 150 RRAGGAPAAEPGELPRPLRRRTARRQPGGRQRTPEAAGGEGAGGRARQEPGQALGRDP 209
QY 50 GSAHVVFSGSGR-AAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPTGLDPL 108
DB 210 GTP---VGAGQGHPAAG-----GLRRLGQPRGLHRRQAHLPPPEP-----AH 247
QY 109 TVRRHVPVWVLLSRDPLDNECGYQPPGAPGLGMPSSCGPRRRRRARDTR 162
DB 248 SAQRQEPF--AAQSRSPRRSAAAG-QPRGTHRA-GIRPAVSEGRQRRLVARPRR 297

Search completed: January 26, 2005, 15:08:51
Job time : 42.4407 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:01 ; Search time 18.2 Seconds  
(without alignments)  
481.084 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_1\_91  
Perfect score: 463  
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPLGLHQLTKLRFQ 91

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137.5	29.7	538	T20H2.2	protein -
2	86.5	18.7	369	2	protein T4012.22 [
3	73.5	15.9	271	2	RHR2 protein - yea
4	73	15.8	907	2	hypothetical prote
5	71.5	15.4	110	2	hypothetical prote
6	70	15.1	367	2	hypothetical prote
7	70	15.1	1160	2	beta3 protein - fr
8	69.5	15.0	257	2	small nuclear ribo
9	69	14.9	639	1	calcium-dependen
10	67.5	14.6	172	2	conserved hypotet
11	67.5	14.6	250	2	hyperosmolarity-re
12	66.5	14.4	467	2	dihydrolipoamide d
13	66.5	14.4	3343	2	breast cancer susc
14	66	14.3	472	1	serine/threonine-s
15	65.5	14.1	89	2	conserved hypotet
16	65.5	14.1	474	2	hypothetical prote
17	65.5	14.1	516	2	probable SF16 prot
18	65	14.0	471	2	hypothetical prote
19	64.5	13.9	482	1	platelet-derived e
20	64.5	13.9	533	1	calcium-dependen
21	64	13.8	127	2	hypothetical prote
22	64	13.8	406	2	hypothetical prote
23	64	13.8	433	2	sensor histidine k
24	64	13.8	455	2	probable carotenoi
25	64	13.8	468	2	probable PPE prote
26	64	13.8	471	2	hypothetical prote
27	64	13.8	750	2	hypothetical prote
28	64	13.8	838	2	hypothetical prote
29	64	13.8	1182	2	myelin transcripti

transcription fact  
zinc finger protei  
hypothetical prote  
probable aldehyde  
hypothetical prote  
hypothetical prote  
heparan sulfate pr  
pyocin negative re  
hypothetical prote  
calcium-dependant  
phospholipase C (E  
non-hemolytic phos  
hypothetical prote  
PU.1-binding prote  
serine/threonine p

#### ALIGNMENTS

##### RESULT 1

H86335  
T20H2.2 protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: H86335

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: H86335

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-538 <STO>

A;Cross-references: UNIPROT:Q9LNUS; GB:AE005172; NID:g8778978; PIDN:AAF79893.1; GSPDB:G

C;Genetics:

A;Map position: 1

Query Match 29.7%; Score 137.5; DB 2; Length 538;  
Best Local Similarity 34.1%; Pred. No. 1.1e-06;  
Matches 30; Conservative 21; Mismatches 28; Indels 9; Gaps 3;  
QY 1 MEHYRKAGSVELP-APSPMPQLPPTLEMRVRDGSKIRLLGLALGRLEGSGARHVVFSG 59  
Db 224 MDKQKQ---VEPKADTPAE-----NEIRITSGRARNVITYMALLQENKSEVIFKA 275  
QY 60 SGRAGKAVSCAEIVKRRVPLGLHQLTKL 87  
Db 276 MGRANKSVTIVELIKRRIPGLHQLTISI 303

##### RESULT 2

F96788

protein T4012.22 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: F96788

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

```

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96788
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-369 <STO>
A;Cross-references: UNIPROT:Q9LQR7; GB:AE005173; NID:g8778814; PIDN:AAF79819.1; GSPDB:GN
C;Genetics:
A;Gene: T4O12.22
A;Map position: 1

Query Match 18.7%; Score 86.5; DB 2; Length 369;
Best Local Similarity 21.4%; Pred.No. 0.17;
Matches 28; Conservative 15; Mismatches 37; Indels 51; Gaps 3;

Qy 1 MEHYRKAGSVELPAPSPMPOLPPDTLEMRVDRGSKIRNLGLALGRLE----- 48
Db 1 MDKQY-----VVKPKATPDIDANEIRITSQGRARNYITYANTLLQVFNSEMCQFYI 53
Qy 49 -----GGSGARHVFGSGGARGAGKAVSCAEIVK-----R 76
Db 54 HCLVFAVCWEIWFDRKSGSTEVFKMGRAINKTIVTVELIKVFFWRPALSNSRSVVQWGR 113
Qy 77 RVPCLHQLTKL 87
Db 114 RIPDLHQNTSI 124

RESULT 3
S48426
RHR2 protein - yeast (Saccharomycetes cerevisiae)
N;Alternate names: protein YIL053w
C;Species: Saccharomycetes cerevisiae
C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: S48426; S63625
R;Smith, V.
submitted to the EMBL Data Library, September 1994
A;Reference number: S48407
A;Accession: S48426
A;Molecule type: DNA
A;Residues: 1-271 <SMI>
A;Cross-references: UNIPROT:P41277; GB:Z47047; EMBL:Z38060; NID:g603997; PID:g763293; MI
R;Hirayama, T.; Maeda, T.; Saico, H.; Shinozaki, K.
Mol. Gen. Genet. 249, 127-138, 1995
A;Title: Cloning and characterization of seven cDNAs for hyperosmolarity-responsive (HOR
A;Reference number: S63623; MUID:96086928; PMID:7500933
A;Accession: S63625
A;Molecule type: mRNA
A;Residues: 17-266 <HIR>
A;Cross-references: EMBL:D50471
C;Genetics:
A;Gene: SGD:RHR2
A;Cross-references: SGD:S0001315; MIPS:YIL053w
A;Map position: 9L

```

[illegible]

RESULT 4  
E96636  
hypothetical protein T7P1.21 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: E96636  
C;Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,  
ansen, N.F.; Hughes, B.; Huizlar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroe, J.S.; Maity, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
kef, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: E96636  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-907 <STO>  
A;Cross-references: UNIPROT:Q9C946; GB:AE005173; NID:g6751696; PIDN:AAF27679.1; GSPDB:GN  
C;Genetics:  
A;Gene: T7p1.21  
A;Map position: 1

	Query Match	15.8%	Score 73;	DB 2;	Length 907;	
	Best Local Similarity	28.4%;	Pred. No. 12;			
	Matches	23; Conservative	13; Mismatches	43; Indels	2; Gaps	1;
Qy	13	PAPSPMPQLPPDTEMRVRDGSKIRNLGLALGRLEGSSARHVVFSGSRAAGKANVSCAE	72			
		: : : : :	: : :		:	:
Dd	639	PPPGAARSLRPKCAATLKRSTGLNLYRILKGKVEGRDPNAKTGSGSRKAGAGSAPAG	698			
Qy	73	IVKRRVPGLHQLTK--LRFLQ	91			
Dd	699	GKQGMDALAETIKKSAYFLQ	719			
	:	:				

RESULT 5  
C72472 hypothetical protein APE2420 - Aeropyrum pernix (strain K1)  
C/Species: Aeropyrum pernix  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C/Accession: C72472  
R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haitawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; ;  
DNA Res. 6, 83-101, 1999  
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyz  
A/Reference number: A72450; MUID:99310339; PMID:10382966  
A/Accession: C72472  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-110 <KAW>  
A/Cross-references: UNIPROT:Q9Y965; DBJ:AF000064; NTD:G5105945; PIDN:BAA81435.1; PID:d1  
A/Experimental source: strain K1  
C/Genetics:  
A/Gene: APE2420

```

Query Match      15.4%; Score 71.5; DB 2; Length 110;
  Best Local Similarity 32.5%; Pred. No. 1.8;
Matches 25; Conservative 10; Mismatches 27; Indels 15; Gaps 4;

Qy 5 RKAGSVELPAPSPMPQ-LPPDTLEMRVRDGSKTRNL-----GIALGRLEGSGARHV--- 56
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 34 REASSSTLKPVGVPKPLPETMTSAFSGTSLTGILTLFTSLAIG--SGGSSTSLTLP 91

Qy 57 -----FSGSGRAAGKAV 69
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 92 CLPGCLSGSGNTLGLAV 108
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 6  
S25185  
hypothetical protein 4 - Streptomyces coelicolor (fragment)  
C:Species: Streptomyces coelicolor  
C:Date: 04-Feb-1998 #sequence revision 20-Feb-1998 #text change

C;Accession: S25185; S21012  
R;Damman, T.; Wohleben, W.  
Mol. Microbiol. 6, 2267-2278, 1992  
A;Title: A metalloprotease gene from *Streptomyces coelicolor* "Mueller" and its transcription  
A;Reference number: S25185, MUID:93023855; PMID:1406267  
A;Accession: S25185  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-367 <DAM>  
A;Cross-references: UNIPROT:P43167; EMBL:Z11929; NID:g46867; PIDN:CAA77983.1; PID:g46867  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1992

Query Match 15.1%; Score 70; DB 2; Length 367;  
Best Local Similarity 27.9%; Pred. No. 9.4;  
Matches 31; Conservative 9; Mismatches 29; Indels 42; Gaps 6;

QY 7 AGSVLPAPSPMPQLPDPDTELMVRDGSKIRNLGL-----AL 44  
Db 167 AARPQLPAP---PQLPPLVPQTPDRPLGRRALGLLHLDRHRAAPRRRAGTLRGLAAL 223  
QY 45 GRLEGGSARHVVFGSGRAAGKAV---SCAEIVK---RRV-----PGL 81  
Db 224 PVLDDG---HALGRGSGTVAYRCAGVHGCEDEVLRHLVRRVADSGCATRGL 271

RESULT 7  
T13713  
beta3 protein - fruit fly (*Drosophila melanogaster*)  
C;Species: *Drosophila melanogaster*  
C;Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: T13713  
R;Zhang, Y.Q.; Broadie, K.S.  
submitted to the EMBL Data Library, October 1998  
A;Description: Dissecting the functions of beta3 of AP3 complex.  
A;Reference number: Z17706  
A;Accession: T13713  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1160 <ZHA>  
A;Cross-references: UNIPROT:O77290; EMBL:AJ011778; NID:e1330417; PID:e1330418; PIDN:CAAC  
C;Genetics:  
A;Gene: beta3  
A;Cross-references: FlyBase:FBgn0003210  
A;Introns: 214/3; 491/1; 899/1

Query Match 15.1%; Score 70; DB 2; Length 1160;  
Best Local Similarity 33.8%; Pred. No. 32;  
Matches 25; Conservative 6; Mismatches 19; Indels 24; Gaps 5;

QY 1 MEHYRK---AGSVLPAPSPMPQLPDPDTELMVRDGSKIRNLGLALGRLEGGSARHVV 57  
Db 659 LSHLNLNPAQYKELPA---FPVIPP-----SSVRNIAGFMQDKLPGEDSP---- 702

QY 58 SG-----SGRAAG 66  
Db 703 SCHSKDKSGREGGK 716

RESULT 8  
T10586  
small nuclear ribonucleoprotein-associated protein homolog F9f13.90 - *Arabidopsis thaliana*  
C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C;Accession: T10586  
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z16991  
A;Accession: T10586  
A;Molecule type: DNA  
A;Residues: 1-257 <BEV>  
A;Cross-references: UNIPROT:Q9SUN5; EMBL:AL080253; GSPDB:GN00062; ATSP:F9f13.90  
A;Experimental source: cultivar Columbia; BAC clone F9f13  
C;Genetics:

A;Gene: ATSP:F9f13.90  
A;Map position: 4  
C;Superfamily: proline-rich protein

Query Match 15.0%; Score 69.5; DB 2; Length 257;  
Best Local Similarity 29.6%; Pred. No. 7.3;  
Matches 24; Conservative 10; Mismatches 26; Indels 21; Gaps 3;

QY 20 QLPPDTELMVRDGSKIRNLGLAL-----GRLEGGSARHVVFGS 60  
Db 50 KLPP-APKKGKINEEDRTLGLVLRGEVISMVTEGPPPEESRAKASAAAVAGPGI 108

QY 61 GRAAGKAVSCAEIVKRRVFL 81  
Db 109 GRAAGRGVPTGLVQAQ-PGL 128

RESULT 9  
T02784  
calcium-dependent protein kinase (EC 2.7.1.1) - maize (strain W64A)  
C;Species: Zea mays (maize)  
C;Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
C;Accession: T02784  
R;Muriillo, I.; Jaeck, E.; Cordero, M.; San Segundo, B.  
submitted to the EMBL Data Library, July 1998  
A;Description: A calcium-dependent protein kinase possibly involved in pathogen defense  
sis-related PRMs Gene.  
A;Reference number: Z14736  
A;Accession: T02784  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-639 <MUR>  
A;Cross-references: UNIPROT:O82107; EMBL:AJ007366; PIDN:CAA07481.1  
A;Experimental source: strain W64A; seed  
C;Function:  
A;Description: probably involved in pathogen defense in maize plants  
C;Superfamily: calcium-dependent protein kinase; calmodulin repeat homology; protein ki  
C;Keywords: ATP; calcium binding; EF hand; phosphotransferase; serine/threonine-specifi  
F;151-411/Domain: protein kinase homology <KIN>  
F;159-167/Region: protein kinase ATP-binding motif  
F;454-486/Domain: calmodulin repeat homology <EF1>  
F;490-522/Domain: calmodulin repeat homology <EF2>  
F;526-558/Domain: calmodulin repeat homology <EF3>  
F;560-592/Domain: calmodulin repeat homology <EF4>  
F;182/Active site: Lys #status Predicted

Query Match 14.9%; Score 69; DB 1; Length 639;  
Best Local Similarity 25.6%; Pred. No. 22;  
Matches 22; Conservative 15; Mismatches 33; Indels 16; Gaps 2;

QY 5 RKAGSVLPAPSPMPQLPDPDTELM-----RVRDGSKIRNLGLALGRLEGSSA 52  
Db 108 KPGGAANASPSRPRPQVRKVSAGLLGLSVLRRTKNTENLKDKYSLGRRLGGQGGFTT 167

QY 53 RHVVFSGSGRAAGKAVSCAEIVKRRV 78  
Db 168 HLCV-----ERATGKELACKSLKRL 189

RESULT 10  
T39026  
conserved hypothetical protein SPAC6C3.02c - fission yeast (*Schizosaccharomyces pombe*)  
C;Species: *Schizosaccharomyces pombe*  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T39026  
R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, February 1996  
A;Reference number: Z21750  
A;Accession: T39026  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-172 <DEV>  
A;Cross-references: UNIPROT:Q10307; EMBL:Z69731; PIDN:CAA93615.1; GSPDB:GN00066; SPDB:S



Query Match 14.3%; Score 66; DB 1; Length 472;  
Best Local Similarity 31.8%; Pred. No. 33;  
Matches 21; Conservative 9; Mismatches 24; Indels 12; Gaps 3;

Qy	7	AGSVLPAPSPMP---QLPPDTLEMRVDSKIRNLGLALGRLEGSARHVFSGSRA 63
		:           :
Db	66	AGTSNPVAIVEKPVDDQLPDVMIEWKIRDE---RN-----ANREDKDMETTVNGSGTE 116

QY 64 AGKAVS 69

Db 117 TGQVIT 122

## RESULT 15

C69494  
conserved hypothetical protein AF1956 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Aug-2004  
C:Accession: C69494  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.P.;  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: C69494  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-89 <KLE>  
C:Cross-references: UNIPROT:O28323; GB:AE000968; GB:AE000782; NID:G2689291; PIDN:AA889293  
C:Superfamily: Archaeal histone-like protein, Alba type

Query Match 14.1%; Score 65.5; DB 2; Length 89;  
Best Local Similarity 41.2%; Pred. No. 6.3;  
Matches 21; Conservative 8; Mismatches 19; Indels 3; Gaps 3;

Qy 33 GSK-IRNLLGLALGRLEGSARHVVFGSGRAAGKAVSCAEIVKRR-VPGL 81  
 Db 9 GNKPMVNYLATLTQLNEG-ADEVVIKARGRAISRAVDVAEIVRNRFMPGV 58

Search completed: January 26, 2005, 15:07:32  
Job time : 19.2 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:40 ; Search time 97.478 Seconds  
(without alignments)  
537.138 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_1\_91

Perfect score: 463  
Sequence: 1 MEHYRKAGSVLPAPSPMPQ.....EIVKRRVPLGLHQLTKLRFQLQ 91

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	463	100.0	163	2	Q8N5L8
2	440	95.0	163	2	Q99JH1
3	439	94.8	163	2	Q6CYX4
4	264.5	57.1	224	2	Q6DGS1
5	224	48.4	199	2	Q91WE3
6	222	47.9	199	2	Q9BUL9
7	218	47.1	199	2	Q9NX88
8	167	36.1	206	2	Q9V9B6
9	166	35.9	215	2	Q7PTH5
10	162	35.0	206	2	Q6NNC7
11	162	35.0	206	2	AAR96155
12	137.5	29.7	315	2	Q944A2
13	137.5	29.7	538	2	Q9LNU5
14	128.5	27.8	350	2	Q93VA8
15	125.5	27.1	485	2	Q9M8Z6
16	119.5	25.8	272	2	Q8H7C6
17	119.5	25.8	732	2	Q7XJ13
18	111	24.0	241	2	Q7RGW6
19	111	24.0	248	2	Q81AX8
20	109.5	23.7	262	2	Q7XT90
21	94	20.3	211	2	Q81DN4
22	90	19.4	202	2	Q7RFX4
23	86.5	18.7	369	2	Q9LQR7
24	86	18.6	160	2	Q9DEP6
25	83	17.9	534	2	Q9ZPM0
26	82	17.7	275	2	Q9AWP6
27	77.5	16.7	206	2	Q7YWL9
28	74	16.0	296	2	Q82DH1
29	74	16.0	321	2	Q6YXK9
30	74	16.0	321	2	BAD05714
31	74	16.0	321	2	BAD05796

32	73.5	15.9	249	1	GPPI_YEAST	P41277	saccharomy
33	73	15.8	907	2	Q9C946	Q9C946	arabidopsis
34	72	15.6	453	2	Q82CD4	Q82CD4	streptomyce
35	72	15.6	458	2	Q72DN2	Q72DN2	desulfovibr
36	72	15.6	458	2	AAS95377	AAS95377	desulfovibr
37	72	15.6	521	2	Q94KH6	Q94KH6	lycopersico
38	71.5	15.4	110	2	Q9Y965	Q9Y965	aeropyrum p
39	71.5	15.4	249	2	Q8LN54	Q8LN54	oryza sativ
40	71.5	15.4	812	2	Q9FBT4	Q9FBT4	streptomyce
41	71	15.3	295	2	Q72B31	Q72B31	desulfovibr
42	71	15.3	295	2	AAS96284	AAS96284	desulfovibr
43	70.5	15.2	216	2	Q6ZHB3	Q6ZHB3	oryza sativ
44	70.5	15.2	216	2	BAD07586	BAD07586	oryza sat
45	70.5	15.2	251	2	Q7VYE7	Q7VYE7	bordetella

ALIGNMENTS

RESULT 1  
Q8N5L8 PRELIMINARY; PRT; 163 AA.  
AC Q8N5L8;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Chromosome 9 open reading frame 23 protein.  
GN Name=C9orf23;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Skin;  
RX MEDLINE=22398257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Schmutz J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalka U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strausberg R.;  
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL; BC032136; AAH32136.1; -  
DR Genew; HGNC:19909; C9orf23;  
DR InterPro; IPR011574; Alba\_DUF78.  
DR ProDom; PD010497; Alba\_DUF78; 1.  
SQ SEQUENCE 163 AA; 17631 MW; 442C8727191A0BCE CRC64;

Query Match 100.0%; Score 463; DB 2; Length 163;

Best Local Similarity 100.0%; Pred. No. 1e-39; 0; Indels 0; Gaps 0;

Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPQPPDTLEMRVDSKIRNLGLALGRLEGSSARHVVFSGS 60  
|||||





```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA The FANTOM Consortium,
RA "The Riken Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yanamoto K., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka I., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Oheato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota Y., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016085; AAH16085.1; -
DR MBL; AK082822; BAC38636.1; -
DR MGD; MGI:2143151; AI851155.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
KW Hypothetical protein.
SQ SEQUENCE 199 AA; 21037 MW; B9DE105BE3293D64 CRC64;

Query Match 48.4%; Score 224; DB 2; Length 199;
Best Local Similarity 47.8%; Pred. No. 5.7e-15;
Matches 47; Conservative 14; Mismatches 28; Indels 10; Gaps 1;

QY 1 MEHYKAGSVELPA-----PSPMPQLPPTDLEMRVDSKIRNLLGLALGRLEGG 50
Db 1 MENFRKVRSEAPAGDGGSGNSGPFADLAPGAVHVRKESKIRNLLAFATASMAQP 60

QY 51 SARHVVFGSGRAAGKAVSCAEIVKRRVPLHQLTKLRF 89
Db 61 ATRAIVFGCGRATTKTVTCAEILKRRLAGLHVQVRLRY 99

RESULT 6
Q9BUL9 ID Q9BUL9 PRELIMINARY; PRT; 199 AA.
AC Q9BUL9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

```

```

DE RPP25 protein (RNase P protein subunit p25).
GN Names=RPP25;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=21998191; PubMed=12003489;
RA Guerrier-Takada C., Eder P.S., Gopalan V., Altman S.;
RT "Purification and characterization of Rpp25, an RNA-binding protein
RT subunit of human ribonuclease P.";
RL RNA 8:290-295 (2002).
DR EMBL; BC002497; AAH02497.1; -
DR EMBL; BC007270; AAH07270.1; -
DR EMBL; AV034074; AAK54443.1; -
DR InterAct; Q9BUL9; -
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 199 AA; 20632 MW; 9AA494F0297B2A81 CRC64;

Query Match 47.9%; Score 222; DB 2; Length 199;
Best Local Similarity 46.5%; Pred. No. 9.1e-15;
Matches 46; Conservative 15; Mismatches 28; Indels 10; Gaps 1;

QY 1 MEHYKAGSVELPA-----PSPMPQLPPTDLEMRVDSKIRNLLGLALGRLEGG 50
Db 1 MENFRKVRSEAPAGCGGSGPFADLAPGAVHVRKESKIRNLLAFATASMAQP 60

QY 51 SARHVVFGSGRAAGKAVSCAEIVKRRVPLHQLTKLRF 89
Db 61 ATRAIVFGCGRATTKTVTCAEILKRRLAGLHVQVRLRY 99

RESULT 7
Q9NX88 ID Q9NX88 PRELIMINARY; PRT; 199 AA.
AC Q9NX88;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein FLJ20374.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,  
 RA Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,  
 RA Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.,  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK000381; BAA91128.1; -;  
 DR InterPro; IPR011574; Alba DUF78.  
 DR ProDom; PD010497; Alba DUF78; 1.  
 SQ SEQUENCE 199 AA; 20660 MW; 855494F0366B2291 CRC64;  
 Query Match 47.1%; Score 218; DB 2; Length 199;  
 Best Local Similarity 45.5%; Pred. No. 2.4e-14;  
 Matches 45; Conservative 15; Mismatches 29; Indels 10; Gaps 1;  
 QY 1 MEHYRKAGSVELPA-----PSPMPQLPDDTLEMRVDRGSKIRNLGLALRLEGG 50  
 DB 1 MEHFRKVRSEAPAGCGAEGGPGFPADLAFGAVHVRKESKIRNLMAFATASNAQP 60  
 QY 51 SARHVFSGSRAAGKAVSCAEIVKRRVPGHLHQLTKLRF 89  
 DB 61 ATRAIVFSGGRVTTKVTCAELKRLAGLHQVTRLRY 99  
 RESULT 8  
 QY9V9B6 PRELIMINARY; PRT; 206 AA.  
 AC QY9V9B6;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG9422-PA (CG9422-pc).  
 GN ORFNames=CG9422;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.H., Blaziej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.B., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupaki M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,  
 RA Svirskas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacleab J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA FLYBASE;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA FLYBASE;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AS003790; AAM70814.1; -;  
 DR IntAct; Q9V9B6; -;  
 DR FlyBase; Pfgn0033092; CG9422.  
 SQ SEQUENCE 206 AA; 23146 MW; 9FCFDF4C3C9BC5D9 CRC64;  
 Query Match 36.1%; Score 167; DB 2; Length 206;  
 Best Local Similarity 41.7%; Pred. No. 4.5e-09;  
 Matches 40; Conservative 16; Mismatches 32; Indels 8; Gaps 2;  
 QY 1 MEHYRKAGSV-----LPASPMQPLPDDTLEMRVDRGSKIRNLGLALRLEGG 53  
 DB 1 MMHYRKAENVKESKSLDFPDCMPKSKQKFLMHVKGKTKVSNVIEFAQALNKGHR 60  
 QY 54 HVVFGSGSRAAGKAVSCAEIVKRRVPGHLHQLTKLRF 89  
 DB 61 CVVWSGGGVTWKISCAEVLKRSHP-LYQVTRMAY 95

```
RESULT 9
Q7PTH5 PRELIMINARY; PRT; 215 AA.
AC Q7PTH5;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGP0000016002 (Fragment).
GN Name=ENSANG000000013513;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008607; EAA03913.2; -.
FT NON_TER 1 1
FT NON_TER 215 215
SQ SEQUENCE 215 AA; 24004 MW; AA72B25FAA8ABFBE CRC64;

Query Match 35.9%; Score 166; DB 2; Length 215;
Best Local Similarity 43.3%; Pred. No. 5.9e-09;
Matches 39; Conservative 16; Mismatches 31; Indels 4; Gaps 2;

QY 1 MEHYRKAGSVLPAPS---PMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSGARHVV 57
DB 6 MMHYRKAENVKELSKDLFFEDCMPSQKDFLMMHVKGKTKVSNVIEFAQALNKGHR 65

QY 58 SSGSRAAGKAVSCAIEVKRRVPGHLQTLKLF 87
DB 66 SSGSGGGVGVKTCIAEIMKRHFE-LHQVTRI 94

RESULT 10
Q6NNC7 PRELIMINARY; PRT; 206 AA.
AC Q6NNC7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE RE65722P.
GN Name=CG9422;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celiniker S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT011363; AAR96155.1; -.
SQ SEQUENCE 206 AA; 23162 MW; 8B9FD01C279BC5D9 CRC64;

Query Match 35.0%; Score 162; DB 2; Length 206;
Best Local Similarity 40.6%; Pred. No. 1.5e-08;
Matches 39; Conservative 16; Mismatches 33; Indels 8; Gaps 2;

QY 1 MEHYRKAGSVLPAPS---LPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSGAR 53
DB 1 MMHYRKAENVKELSKDLFFEDCMPSQKDFLMMHVKGKTKVSNVIEFAQALNKGHR 60

QY 54 HVVFGSGRAAGKAVSCAIEVKRRVPGHLQTLKLF 89
DB 61 CVVWSGGGVGVKTCIAEVLKRSH-P-LYQVTIMAY 95

RESULT 11
AAR96155 PRELIMINARY; PRT; 206 AA.
ID AAR96155
AC AAR96155;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE RE65722P.
GN CG9422.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celiniker S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT011363; AAR96155.1; -.
SQ SEQUENCE 206 AA; 23162 MW; 8B9FD01C279BC5D9 CRC64;

Query Match 35.0%; Score 162; DB 2; Length 206;
Best Local Similarity 40.6%; Pred. No. 1.5e-08;
Matches 39; Conservative 16; Mismatches 33; Indels 8; Gaps 2;

QY 1 MEHYRKAGSVLPAPS---LPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSGAR 53
DB 1 MMHYRKAENVKELSKDLFFEDCMPSQKDFLMMHVKGKTKVSNVIEFAQALNKGHR 60

QY 54 HVVFGSGRAAGKAVSCAIEVKRRVPGHLQTLKLF 89
DB 61 CVVWSGGGVGVKTCIAEVLKRSH-P-LYQVTIMAY 95

RESULT 12
Q944A2 PRELIMINARY; PRT; 315 AA.
ID Q944A2
AC Q944A2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE At1G20220/T20H2.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosida II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narueaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Shinozaki K., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF439833; AAL27504.1; -.
DR HSP; 028323; INFJ.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 315 AA; 33763 MW; 6DC486E51238A937 CRC64;

Query Match 29.7%; Score 137.5; DB 2; Length 315;
Best Local Similarity 34.1%; Pred. No. 7.7e-06;
Matches 30; Conservative 21; Mismatches 28; Indels 9; Gaps 3;

QY 1 MEHYRKAGSVLPAPS---LPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSGARHVV 53
DB 1 MMHYRKAENVKELSKDLFFEDCMPSQKDFLMMHVKGKTKVSNVIEFAQALNKGHR 60

QY 54 HVVFGSGRAAGKAVSCAIEVKRRVPGHLQTLKLF 89
DB 61 CVVWSGGGVGVKTCIAEVLKRSH-P-LYQVTIMAY 95
```

```

Db 1 MDKYQR---VEKPKADTPIAE-----NEIRITSGRARNVITYAMALLQENKSNVIFKA 52
Qy 60 SGRAAGKAVSCAEIVKRVFGLHQLTKL 87
Db 53 MGRAINKSVTIVELIKRRIPDLHQHTSI 80

RESULT 13
Q9LNUS ID Q9LNUS PRELIMINARY; PRT; 538 AA.
AC Q9LNUS;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE T20H2.2 protein.
GN Name=T20H2.2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Sakano H., Vaysberg M., Lee J.M., Lenz C., Liu S., Pham P.,
RA Toriumi M., Yu G., Chin C., Chioi J., Choi E., Chung M., Gonzalez A.,
RA Howng B., Liu A., Altati H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC022472; AA079893.1; -.
DR PIR; H86335; H86335.
DR HSP; O28323; 1NFJ.
DR InterPro; IPR011574; Alba DUF78.
DR ProDom; PD010497; Alba DUF78; 1.
SQ SEQUENCE 538 AA; 58572 MW; 1B0BB141F0F9AF22 CRC64;

Query Match 29.7%; Score 137.5; DB 2; Length 538;
Best Local Similarity 34.1%; Pred. No. 1.3e-05;
Matches 30; Conservative 21; Mismatches 28; Indels 9; Gaps 3;

Qy 1 MEHYRKAGSVELP-APSPMPQLPDPDTLEMRVDRGSKTRNLLGLALGRLEGSAHVVFSG 59
Db 224 MDKYQR---VEKPKADTPIAE-----NEIRITSGRARNVITYAMALLQENKSNVIFKA 275
Qy 60 SGRAAGKAVSCAEIVKRVFGLHQLTKL 87
Db 276 MGRAINKSVTIVELIKRRIPDLHQHTSI 303

RESULT 14
Q93VA8 ID Q93VA8 PRELIMINARY; PRT; 350 AA.
AC Q93VA8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Atlg76010/T4012.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

```

```

RA Ecker J.R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF428441; AAL16210.1; -.
DR EMBL; AY054208; AAL06869.1; -.
DR EMBL; AF412102; AAL06555.1; -.
DR EMBL; AY124847; AAM70556.1; -.
DR HSP; P74761; 1H0X.
DR InterPro; IPR011574; Alba DUF78.
DR ProDom; PD010497; Alba DUF78; 1.
SQ SEQUENCE 350 AA; 37383 MW; DB474865DF4E162E CRC64;

Query Match 27.8%; Score 128.5; DB 2; Length 350;
Best Local Similarity 32.2%; Pred. No. 7.3e-05;
Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;

Qy 1 MEHYRKAGSVELPAPSPMPQLPDPDTLEMRVDRGSKTRNLLGLALGRLEGSAHVVFSGS 60
Db 1 MDKYQR-----VVKPKADTPIDANEIRITSGRARNVITYAMTLLQDKGSTEVVFKAM 53

RESULT 15
Q9M8Z6 ID Q9M8Z6 PRELIMINARY; PRT; 485 AA.
AC Q9M8Z6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F17A9.19 protein.
GN Name=F17A9.19;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B., Wu D.,

```

RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,  
 RL Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC016827; AAF27007.1; -  
 DR HSP; P74761; IH0X.  
 DR InterPro; IPR011574; Alba\_DUF78.  
 DR ProDom; PD010497; Alba\_DUF78; 1.  
 SQ SEQUENCE 485 AA; 53015 MW; 2CF700B58280F3E0 CRC64;

Query Match 27.1%; Score 125.5; DB 2; Length 485;  
 Best Local Similarity 31.3%; Pred.No. 0.00021;  
 Matches 26; Conservative 17; Mismatches 33; Indels 7; Gaps 1;

Qy 1 MEHYRKAGSVLEPAPSPMPQLPPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60

Db 1 MDRYQVRA-----KPKPSPINENEIRITSKGLIRNYISYATSLLOEKSVDIVLKAM 53

Qy 61 GRAAGKAVSCAEIVKRRVPGHLQ 83

Db 54 GOAISKTVAISEILKNKIPGLHQ 76

Search completed: January 26, 2005, 15:06:27  
 Job time : 99.478 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:04 ; Search time 90.6915 Seconds  
(without alignments)  
359.949 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_1\_91

Perfect score: 463  
Sequence: 1 MEHYRAGSVLPAPSPMPQ.....EIVKRRVPGHLQTLKRLFLQ 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	463	100.0	163	4	AAG89339 Human sec
2	463	100.0	163	6	ABU00014 Human nov
3	463	100.0	219	8	ADM90988 Human pha
4	457	98.7	163	2	RAY36104 Extended
5	457	98.7	163	8	ADP19412 Human sec
6	402	86.8	79	2	RAY11950 Human 5'
7	395	85.3	78	3	AAG03599 Human sec
8	218	47.1	211	4	AM338973 Human pol
9	164	35.4	205	4	ABB64410 Drosophil
10	134.5	29.0	99	3	AAG36977 Arabidops
11	128.5	27.8	350	3	AAG16838 Arabidops
12	128.5	27.8	350	3	AAG51875 Arabidops
13	128.5	27.8	350	3	AAG51847 Arabidops
14	128.5	27.8	350	8	ADN72883 Thale cre
15	128.5	27.8	350	8	ADN72183 Thale cre
16	128.5	27.8	356	3	AAG51846 Arabidops
17	119	25.7	74	3	AAG36978 Arabidops
18	116	25.1	267	3	AAG33136 Arabidops
19	104.5	22.6	259	3	AAG34072 Zea mays
20	104.5	22.6	313	3	AAG34071 Zea mays
21	104	22.5	239	3	AAG33137 Zea mays
22	94	20.3	314	3	AAG51848 Arabidops
23	94	20.3	314	3	AAG16839 Arabidops
24	94	20.3	314	3	AAG51876 Arabidops
25	78	16.8	298	3	AAG51877 Arabidops

26	78	16.8	298	3	AAG16840 Arabidops
27	76	16.4	247	2	AAY48525 Human bre
28	75.5	16.3	255	6	ABU32060 Protein e
29	75.5	16.3	261	7	ABO65891 Klebsiell
30	75.5	16.3	494	5	ABP62768 Protein f
31	75.5	16.3	494	7	ADJ72179 Streptomy
32	73.5	15.9	121	6	ABU00334 Human nov
33	73.5	15.9	271	2	AAM57327 Mitochond
34	73.5	15.9	271	2	AAW30683 Glycerol-
35	73.5	15.9	271	2	AAW60263 Klebsiell
36	73.5	15.9	271	2	AAY26169 Cycosollic
37	72.5	15.7	290	7	ABO74266 Pseudomon
38	72	15.6	164	3	AAY44836 Human Per
39	72	15.6	213	6	ADA54493 Human pro
40	72	15.6	215	3	AAG33138 Zea mays
41	71	15.3	240	5	ADK35009 Novel hum
42	71	15.3	566	7	ABO69785 Pseudomon
43	70.5	15.2	655	4	ABG23459 Novel hum
44	70.5	15.2	1203	7	ADJ70368 Human hea
45	70	15.1	246	3	AAG48477 Arabidops

ALIGNMENTS

RESULT 1  
AAG89339  
ID AAG89339 standard; protein; 163 AA.  
XX AC AAG89339;  
XX AC  
DT 11-SEP-2001 (first entry)  
XX  
DE Human secreted protein, SEQ ID NO: 459.  
XX  
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
GENSET.  
XX Homo sapiens.  
XX  
FN WO200142451-A2.  
XX  
PD 14-JUN-2001.  
XX  
PF 07-DEC-2000; 2000WO-IB001938.  
XX  
PR 08-DEC-1999; 99US-0169629P.  
PR 06-MAR-2000; 2000US-0187470P.  
(GEST ) GENSET.  
Dumas Milne Edwards J, Bougueleret L, Jobert S;  
WPI; 2001-367870/38.  
N-PSDB; AAH64942.  
Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.  
Claim 21; Page 910-911; 921pp; English.  
The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar nucleic acid sequences in samples, and hence to determine which  
 CC patients may be in need of restorative therapy. The GENSER polypeptides  
 CC may be used as antigens in the production of antibodies and in assays to  
 CC identify modulators (agonists and antagonists) of GENSER polypeptide  
 CC expression and activity. The present sequence is a GENSER polypeptide of  
 CC the invention  
 XX the invention  
 XX

SQ Sequence 163 AA;

Query Match 100.0%; Score 463; DB 4; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-47;  
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60  
 |||||  
 DB 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60  
 |||||  
 QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
 |||||  
 DB 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
 |||||

RESULT 2

ABU00014  
 ID ABU00014 standard; protein; 163 AA.

XX AC ABU00014;

XX DT 17-JAN-2003 (first entry)

XX DE Human novel polypeptide #107.

XX KW Human; genetic disorder; gene mapping; medical imaging; cancer;  
 KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;  
 KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;  
 KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;  
 KW fungal infection; bacterial infection; autoimmune disease; diabetes;  
 KW atopic dermatitis.

XX OS Homo sapiens.

XX XN WO200274961-A1.

XX PD 26-SEP-2002.

XX PF 14-MAR-2002; 2002WO-US0005109.

XX PR 15-MAR-2001; 2001US-00810173.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX DR WPI; 2003-040556/03.  
 XX DR N-PSDB; ABX05092.

XX PT New isolated polypeptides and polynucleotides, useful for preventing,  
 PT treating or ameliorating medical conditions, such as cancer,  
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative  
 PT disorders, and infections.

XX PS Claim 9; SEQ ID NO 633; 235pp; English.

XX CC The invention relates to human polynucleotides and the polypeptides they  
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,  
 CC forensics, gene mapping, medical imaging, identification of mutations  
 CC responsible for genetic disorders or other traits, assessing biodiversity  
 CC and producing many other types of data and products dependent on DNA and  
 CC amino acid sequences. They are also useful for preventing, treating or  
 CC ameliorating medical conditions, such as cancer, neurodegenerative

CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell  
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,  
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or  
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).  
 CC Sequences ABG9988-ABG9989 and ABU0010-ABU00433 represent human  
 CC polypeptides of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied by the European Patent Office  
 XX

SQ Sequence 163 AA;

Query Match 100.0%; Score 463; DB 6; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-47;  
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60  
 |||||  
 DB 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60  
 |||||  
 QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
 |||||  
 DB 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
 |||||

RESULT 3

ADM90988

ID ADM90988 standard; protein; 219 AA.

XX AC ADM90988;

XX DT 03-JUN-2004 (first entry)

XX DE Human pharmaceutically useful protein SeqID 381.

XX KW human; cancer; haematopoiesis; thrombosis; anaemia;  
 KW cardiovascular disorder; ischaemic heart disease;  
 KW acute myocardial infarction; respiratory disease; asthma; pneumonia;  
 KW cystic fibrosis; chronic renal failure; glomerulopathy;  
 KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;  
 KW HIV infection; systemic lupus erythematosus; endocrine system;  
 KW diabetes mellitus; epilepsy; Alzheimer's disease;  
 KW amytropic lateral sclerosis; skin disorder; psoriasis; bacterial;  
 KW fungal; parasitic; viral infection; cytostatic; anticoagulant;  
 KW thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;  
 KW antiinflammatory; nephrotropic; antitumor; hepatotropic;  
 KW immunosuppressive; antiallergic; dermatological; antirheumatic;  
 KW antiarthritic; antidiabetic; anticonvulsant; neuroprotective; nootropic;  
 KW antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;  
 KW gene therapy; vaccine.

XX OS Homo sapiens.

XX XN WO2004020595-A2.

XX PD 11-MAR-2004.

XX PF 28-AUG-2003; 2003WO-US027107.

XX PR 29-AUG-2002; 2002US-0406576P.

XX PR 29-AUG-2002; 2002US-0406611P.

XX PR 29-AUG-2002; 2002US-0406612P.

XX PR 29-AUG-2002; 2002US-0406616P.

XX PR 29-AUG-2002; 2002US-0406640P.

XX PR 29-AUG-2002; 2002US-0406655P.

XX PR 17-SEP-2002; 2002US-0406666P.

XX PR 17-SEP-2002; 2002US-0410946P.

XX PR 17-SEP-2002; 2002US-0410951P.

XX PR 17-SEP-2002; 2002US-0410953P.

XX PR 17-SEP-2002; 2002US-0410957P.

XX PR 17-SEP-2002; 2002US-0410960P.

XX PR 17-SEP-2002; 2002US-0410962P.

XX PR 17-SEP-2002; 2002US-0411019P.

XX PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.  
 PA (DNAF-) DNAFORM KK.  
 XX Williams LT, Chu K, Lee E, Hestir K;  
 XX  
 DR WPI; 2004-257410/24.  
 DR N-PSDB; ADM90779, ADM91197.  
 XX  
 XX New human polynucleotides and polypeptides, useful for diagnosing,  
 PT preventing and treating proliferative disorders, immune disorders,  
 PT cardiovascular disorders, or bacterial, fungal, parasitic and viral  
 PT diseases.  
 XX  
 XX Claim 1; SEQ ID NO 381; 254bp; English.  
 XX  
 CC This invention relates to novel isolated human polynucleotides and the  
 CC encoded proteins thereof. Specifically, it refers to proteases, kinases,  
 CC phosphatases, secreted and transmembrane proteins, as well as the derived  
 CC peptide fragments, which can be used to develop antibodies and screen for  
 CC small molecule agonists and antagonists that can modulate their  
 CC activities. The present invention describes polypeptides,  
 CC polynucleotides, vectors and host cells useful for diagnosing, preventing  
 CC and treating proliferative disorders, e.g. cancer, disorders of  
 CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,  
 CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory  
 CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the  
 CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,  
 CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune  
 CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders  
 CC of the endocrine system, e.g. diabetes mellitus, central nervous system  
 CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral  
 CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,  
 CC parasitic and viral diseases. Accordingly, they exhibit many various  
 CC activities including cytostatic, anticoagulant, thrombolytic,  
 CC antinaeumic, cardiac, vasotropic, antidiabetic, antiinflammatory,  
 CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,  
 CC dermatological, antirheumatic, antiarthritic, antidiabetic,  
 CC anticonvulsant, neuroprotective, nontropic, antipsoriatic, antibacterial,  
 CC fungicide, antiparasitic and virucidal, such that these polynucleotides  
 CC can be used for gene therapy purposes and the development of appropriate  
 CC vaccines. This polypeptide is a human protein of the invention.  
 XX  
 XX Sequence 219 AA;  
 SQ  
 Query Match 100.0%; Score 463; DB 8; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-46;  
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEHYRKAGSVLPAPSPMPQPDPTLEMRVRDGSKIRNLLGLALGLEGGSARHVVFGS 60  
 Db 57 MEHYRKAGSVLPAPSPMPQPDPTLEMRVRDGSKIRNLLGLALGLEGGSARHVVFGS 116  
 QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
 Db 117 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 147  
 RESULT 4  
 ID AAY36104  
 XX  
 AC AAY36104 standard; protein; 163 AA.  
 XX  
 AC AAY36104;  
 XX  
 DT 13-SEP-1999 (first entry)  
 XX  
 XX

DE Extended human secreted protein sequence, SEQ ID NO. 489.  
 XX  
 KW Secreted protein; human; cytokine; cellular proliferation; cell movement;  
 KW cellular differentiation; immune system regulator; anti-inflammatory;  
 KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;  
 KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;  
 KW genetic disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9931236-A2.  
 XX  
 PD 24-JUN-1999.  
 XX  
 PF 17-DEC-1998; 98WO-IB002122.  
 XX  
 PR 17-DEC-1997; 97US-0069957P.  
 PR 09-FEB-1998; 98US-0074121P.  
 PR 13-APR-1998; 98US-0081563P.  
 PR 10-AUG-1998; 98US-0096116P.  
 XX  
 XX (GEST ) GENSET.  
 PA  
 XX  
 XX Bougueleret L, Duclert A, Dumas Milne Edwards J;  
 PI  
 XX WPI; 1999-385906/32.  
 DR N-PSDB; AAX97788.  
 DR  
 XX  
 XX New isolated human secreted proteins.  
 PT  
 XX Claim 9; Page 423-424; 516pp; English.  
 PS  
 XX This sequence is encoded by an extended human secreted protein coding  
 CC sequence of the invention. The secreted proteins can be used in treating  
 CC or controlling a variety of human conditions. The secreted proteins may  
 CC act as cytokines or may affect cellular proliferation or differentiation  
 CC or may act as immune system regulators, haematopoiesis regulators, tissue  
 CC growth regulators, regulators of reproductive hormones or cell movement  
 CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or  
 CC tumour inhibition activity. The DNAs can be used in forensic procedures  
 CC to identify individuals or in diagnostic procedures to identify  
 CC individuals having genetic diseases resulting from abnormal expression of  
 CC the genes corresponding to the extended cDNAs. They are also useful for  
 CC constructing a high resolution map of the human chromosomes. They can  
 CC also be used for gene therapy to control or treat genetic diseases  
 XX  
 XX Sequence 163 AA;  
 SQ  
 Query Match 98.7%; Score 457; DB 2; Length 163;  
 Best Local Similarity 98.9%; Pred. No. 3.9e-46;  
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MEHYRKAGSVLPAPSPMPQPDPTLEMRVRDGSKIRNLLGLALGLEGGSARHVVFGS 60  
 Db 1 MEHYRKAGSVLPAPSPMPQPDPTLEMRVRDGSKIRNLLGLALGLEGGSARHVVFGS 60  
 QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
 Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
 RESULT 5  
 ID ADP19412  
 XX ADP19412 standard; protein; 163 AA.  
 AC ADP19412;  
 XX  
 DT 26-AUG-2004 (first entry)  
 XX  
 DE Human secreted polypeptide #263.  
 XX  
 XX Human; secreted protein; genetic disease.  
 XX

OS Homo sapiens.  
 XX US2004110939-A1.  
 XX 10-JUN-2004.  
 XX 15-OCT-2001; 2001US-00978360.  
 XX 17-DEC-1998; 98WO-IB002122.  
 XX 09-FEB-1999; 99WO-IB000282.  
 XX 21-JUN-2000; 2000WO-IB000951.  
 XX 15-SEP-2000; 2000US-00663600.  
 XX (GEST ) GENSET SA.  
 XX Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;  
 XX Duclert A;  
 XX WPI; 2004-440404/41.  
 XX N-PSDB; ADP19007.  
 XX New isolated polynucleotide encoding secreted polypeptide, useful for  
 PT gene therapy, or in diagnostic procedures to identify individuals having  
 PT genetic diseases resulting from abnormal expression of the genes.  
 XX Claim 2; SEQ ID NO 668; 113pp; English.  
 XX The invention relates to human cDNA sequences that encode human secreted  
 CC proteins. The invention also relates to an antibody that specifically  
 CC binds to a polypeptide of the invention and a method of binding the  
 CC polypeptide to an antibody. The polynucleotides are useful for expressing  
 CC the entire secreted proteins which they encode and for distinguishing  
 CC human tissues and cells from non-human tissues and cells, and for  
 CC distinguishing between human tissues and cells that do or do not express  
 CC the polynucleotides comprising the cDNAs. The polynucleotides and  
 CC polypeptides are useful in forensic procedures or diagnostic procedures  
 CC to identify individuals with genetic diseases resulting from abnormal  
 CC expression of the genes corresponding to the cDNAs. The sequences are  
 CC also useful in gene therapy to control or treat genetic diseases. This  
 CC sequence represents a human secreted polypeptide of the invention. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html.  
 XX seqdata.uspto.gov/sequence.html.  
 XX Sequence 163 AA;  
 SQ  
 Query Match 98.7%; Score 457; DB 8; Length 163;  
 Best Local Similarity 98.9%; Pred. No. 3.9e-46;  
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGS 60  
 DB 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGS 60  
 OY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLFLQ 91  
 DB 61 GRAAGKAVSCAEIVKRRVPGHQLTKLFLQ 91  
 RESULT 6  
 AAY11950  
 ID AAY11950 standard; protein; 79 AA.  
 XX AAY11950;  
 XX 18-JUN-1999 (first entry)  
 XX Human 5' EST secreted protein SEQ ID No: 550.  
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;  
 KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;  
 KW upstream regulatory sequence; cytokine activity; cell proliferation;  
 KW differentiation; haematopoiesis regulation; tissue growth regulation;  
 KW  
 KW reproductively hormone regulation; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; anti-inflammatory; tumour inhibition.  
 XX Homo sapiens.  
 XX WO9906550-A2.  
 XX 11-FEB-1999.  
 XX 31-JUL-1998; 98WO-IB001232.  
 XX 01-AUG-1997; 97US-00905144.  
 XX (GEST ) GENSET.  
 XX Dumas Milne Edwards J, Duclert A, Lacroix B;  
 XX WPI; 1999-153780/13.  
 XX N-PSDB; AAX40672.  
 XX New isolated prostate-derived nucleic acids - used to develop products  
 PT which may have cytokine, immune regulatory, haematopoiesis regulating,  
 PT anti-inflammatory or tumour inhibition activity.  
 XX Claim 34; Page 648; 675pp; English.  
 XX AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for  
 CC human secreted proteins expressed in prostate, and encode the proteins  
 CC given in AAY11716 to AAY11993 respectively. The proteins given represent  
 CC the signal peptide and an N-terminal fragment of a secreted protein. The  
 CC nucleic acid sequences can be used for producing secreted human gene  
 CC products. They can also be used to develop products for diagnosis and  
 CC therapy. The proteins obtained may have cytokine activity, cell  
 CC proliferation and differentiation activity, haematopoiesis regulating  
 CC activity, tissue growth regulating activity, reproductively hormone  
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and  
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter  
 CC sequences. The nucleic acids encoding the signal peptides can be used for  
 CC directing extracellular secretion of a polypeptide or the insertion of a  
 CC polypeptide into a membrane, or importing a polypeptide into a cell  
 XX  
 XX Sequence 79 AA;  
 SQ  
 Query Match 86.8%; Score 402; DB 2; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-40;  
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGS 60  
 DB 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGS 60  
 OY 61 GRAAGKAVSCAEIVKRRVP 79  
 DB 61 GRAAGKAVSCAEIVKRRVP 79  
 RESULT 7  
 AAG03599  
 ID AAG03599 standard; protein; 78 AA.  
 XX AAG03599;  
 XX 06-OCT-2000 (first entry)  
 XX Human secreted protein, SEQ ID NO: 7680.  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.  
 XX Homo sapiens.

```

XX PN EF1033401-A2.
XX PD
XX PF
XX PP
XX PR 21-FEB-2000; 2000EP-00200610.
XX PR 26-FEB-1999; 99US-0122487P.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX DR N-PSDB; AAC03605.
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX PS Claim 13; SEQ ID NO 7680; 71pp + Sequence Listing; English.
XX CC The present sequence is a polypeptide encoded by one of a large number of
XX CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
XX CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
XX CC tissues. EST sequences usually correspond mainly to the 3' untranslated
XX CC region (UTR) of the mRNA because they are often obtained from oligo-dT
XX CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
XX CC sequences derived from the 5' ends of mRNAs and even in those cases where
XX CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
XX CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
XX CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
XX CC are also used in diagnostic, forensic, gene therapy and chromosome
XX CC mapping procedures. They are used to obtain upstream regulatory sequences
XX CC and to design expression and secretion vectors
XX SQ Sequence 78 AA;
Query Match 85.3%; Score 395; DB 3; Length 78;
Best Local Similarity 100.0%; Pred. No. 4e-39;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVELPAPSPMPQLPDTLEMRVRDGSKIRNLGLALGRLEGGSARHVFSGS 60
DB 1 MEHYRKAGSVELPAPSPMPQLPDTLEMRVRDGSKIRNLGLALGRLEGGSARHVFSGS 60
QY 61 GRAAGKAVSCAEIVKRRV 78
DB 61 GRAAGKAVSCAEIVKRRV 78
RESULT 8
AM38973
ID AA38973 standard; protein; 211 AA.
AC AA38973;
XX AA38973;
XX 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 2118.
XX Human; neotrophic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.

```

```

XX XX 26-DEC-2000; 2000WO-US034263.
XX PF
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-0048725.
XX PR 25-APR-2000; 2000US-0052317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX DR N-PSDB; AA158129.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX Example 4; SEQ ID NO 2118; 10078pp; English.
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AA158642-AA162213) with neotrophic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX SQ Sequence 211 AA;
Query Match 47.1%; Score 218; DB 4; Length 211;
Best Local Similarity 45.5%; Pred. No. 1.7e-17;
Matches 45; Conservative 15; Mismatches 29; Indels 10; Gaps 1;
QY 1 MEHYRKAGSVEL-----PAPSPMPQLPDTLEMRVRDGSKIRNLGLALGRLEG 50
DB 13 MENFRKVRSEEAAPRGCGAEGGPGFPADLAPGAVHMRVKEGSKIRNLMAFATASMAQP 72
QY 51 SARHVVFGSGRAAGKAVSCAEIVKRRVFGHLQTLKRF 89
DB 73 ATRAI VFSGCGRATTKTTCABEILKERLAGLHQVTRLY 111
RESULT 9
ABB64410
ID ABB64410 standard; protein; 205 AA.
XX ABB64410;
XX ABB64410;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 20022.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX OS
XX XX

```



```
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144804P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145119P.
PR 23-JUL-1999; 99US-0145115P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147119P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 08-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 16-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 03-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.

PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 29.0%; Score 134.5; DB 3; Length 99;
Best Local Similarity 34.8%; Pred. No. 6.1e-08;
Matches 30; Conservative 20; Mismatches 29; Indels 9; Gaps 3;

QY 1 MEHYRKAGSVLEP-APSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSG 59
Db 1 MDKYQR--VEKEKADTPAE-----NEIRITSMGRARVITYAMALLXENKSNVIFKA 52

QY 60 SGRAAGKAVSCAIVKRRVPGHLHQLTKL 87
Db 53 MGRAINKSVTIVELIKRRIPGLHQITSI 80

RESULT 11
AAG16838
ID AAG16838 standard; protein; 350 AA.
XX
AC AAG16838;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17635.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
```

PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131144P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132248P.  
PR 05-MAY-1999; 99US-0132248P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142220P.  
PR 12-JUL-1999; 99US-0142377P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149923P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151068P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.



```
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161930P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 27.8; Score 128.5; DB 3; Length 350;
Best Local Similarity 32.2; Pred. No. 1.4e-06;
Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;

Qy 1 MEHYKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLGLALGRLEGGSARHVPSGS 60
Db 1 MDKQK-----VVKPATPIDANEIRITSGRARNYTYAWTLQLDKGSTEVVFKAM 53

Qy 61 GRAAGKAVSCAEIVKRVPGHLHQLTKL 87
Db 54 GRAINKTVTIVELIKRIPDLHQNTSI 80

RESULT 12
AAG51875
ID AAG51875 standard; protein; 350 AA.
XX
AC AAG51875;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65883.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123160P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.

16-APR-1999; 99US-0129845P.
19-APR-1999; 99US-0130077P.
21-APR-1999; 99US-0130449P.
23-APR-1999; 99US-0130510P.
23-APR-1999; 99US-0130891P.
28-APR-1999; 99US-0131449P.
30-APR-1999; 99US-0132048P.
04-MAY-1999; 99US-0132407P.
05-MAY-1999; 99US-0132484P.
06-MAY-1999; 99US-0132485P.
06-MAY-1999; 99US-0132486P.
06-MAY-1999; 99US-0132487P.
07-MAY-1999; 99US-0132863P.
11-MAY-1999; 99US-0134256P.
14-MAY-1999; 99US-0134218P.
14-MAY-1999; 99US-0134219P.
14-MAY-1999; 99US-0134221P.
14-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
20-MAY-1999; 99US-0135124P.
21-MAY-1999; 99US-0135353P.
24-MAY-1999; 99US-0135629P.
25-MAY-1999; 99US-0136021P.
27-MAY-1999; 99US-0136392P.
28-MAY-1999; 99US-0136782P.
01-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
```

PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145122P.  
PR 23-JUL-1999; 99US-0145122P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0148171P.  
PR 06-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 21-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.

PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160880P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 27.8%; Score 128.5; DB 3; Length 350;  
Best Local Similarity 32.2%; Pred. No. 1.4e-06;  
Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;

QY 1 MEHYKAGSVLPAPSPMPQLPPDPTLEMRVDRGSKIRNLGLALGRLEGGARHVVFSGS 60  
DB 1 MDKYOR-----VVKRADTPIDANEIRITSQGRARNYITYAMTLLQDKGSTEVVFKAM 53  
QY 61 GRAGKAVSCAEIVKRRVPEGLHQLTKL 87  
DB 54 GRAINKTIVIVELIKRIPDLHONTSI 80

## RESULT 13

AAG51847  
ID AAG51847 standard; protein; 350 AA.

XX AC AAG51847;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 65844.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX OS termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128214P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 08-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 19-MAY-1999; 99US-0134768P.  
PR 20-MAY-1999; 99US-0134941P.  
PR 21-MAY-1999; 99US-01351124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136031P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 08-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 09-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 08-OCT-1999; 99US-0158029P.  
PR 12-OCT-1999; 99US-0158232P.  
PR 13-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160880P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 28-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 27.8%; Score 128.5; DB 3; Length 350;  
 Best Local Similarity 32.2%; Pred. No. 1.4e-06;

Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;

QY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60

DB 1 MDKYQR-----VVKPKADTPIDANEIRITSGRARNYITYAMTLLQDKGSTEVVFKAM 53

QY 61 GRAAGKAVSCAEIVKRVPGHQLTKL 87

DB 54 GRAINKTIVTIVELIKRIPDLHQNTSI 80

# RESULT 14

ID ADN72883  
 ID ADN72883 standard; protein; 350 AA.

XX AC ADN72883;

XX DT 15-JUL-2004 (first entry)

XX DE Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 778.

XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;

XX KW animal feed product; thale cress; cell wall biosynthesis;

XX KW nitrogen metabolism; carbon metabolism.

XX OS Arabidopsis thaliana.

XX PN WO2004035798-A2.

XX PD 29-APR-2004..

XX PF 20-OCT-2003; 2003WO-EP011658.

XX PR 18-OCT-2002; 2002EP-00079408.

XX PA (CROP-) CROPDESIGN NV.

XX PI Inze D, De Veylder L, Vlieghe K;

XX DR WPI; 2004-348466/32.

XX DR N-PSDB; ADN72882.

XX PT Altering plant characteristics, useful for producing plants for enzyme or  
 PT pharmaceutical production comprises modifying in a plant, expression of  
 PT one or more nucleic acids and/or modifying level or activity of one or  
 PT more proteins.

XX PS Claim 1; SEQ ID NO 778; 134pp; English.

XX

CC This invention relates to a novel method for altering one or more plant  
 CC characteristics. Specifically, it refers to identifying genes that are up  
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
 CC alter plant characteristics accordingly. The present invention describes  
 CC generating transgenic plants for the production of growth regulators,  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
 CC the altered plant characteristics are selected from increased yield or  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
 CC or physiology, altered endoreplication, biochemistry, signal  
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,  
 CC each relative to the corresponding wild type plants. Accordingly, these  
 CC sequences can also be useful as positive or negative selectable markers  
 CC during transformation of cells or tissues. The identified genes play a  
 CC role in a variety of biological processes such as DNA replication, cell  
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
 CC transcription factors. This polypeptide sequence is thale cress protein  
 CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing  
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the  
 CC invention.

XX SQ Sequence 350 AA;

Query Match 27.8%; Score 128.5; DB 8; Length 350;

Best Local Similarity 32.2%; Pred. No. 1.4e-06;

Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;

QY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60

DB 1 MDKYQR-----VVKPKADTPIDANEIRITSGRARNYITYAMTLLQDKGSTEVVFKAM 53

QY 61 GRAAGKAVSCAEIVKRVPGHQLTKL 87

DB 54 GRAINKTIVTIVELIKRIPDLHQNTSI 80

# RESULT 15

ADN72183

ID ADN72183 standard; protein; 350 AA.

XX AC ADN72183;

XX DT 15-JUL-2004 (first entry)

XX DE Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 78.

XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;

XX KW animal feed product; thale cress; cell wall biosynthesis;

XX KW nitrogen metabolism; carbon metabolism.

XX OS Arabidopsis thaliana.

XX PN WO2004035798-A2.

XX PD 29-APR-2004.

XX PF 20-OCT-2003; 2003WO-EP011658.

XX PR 18-OCT-2002; 2002EP-00079408.

XX PA (CROP-) CROPDESIGN NV.

XX PI Inze D, De Veylder L, Vlieghe K;

XX DR WPI; 2004-348466/32.

XX DR N-PSDB; ADN72182.

XX PT Altering plant characteristics, useful for producing plants for enzyme or  
 PT pharmaceutical production comprises modifying in a plant, expression of  
 PT one or more nucleic acids and/or modifying level or activity of one or  
 PT more proteins.



**This Page Blank (uspto)**

Result No.	Query			Description	
	Score	Match	Length	ID	Description
1	463	100.0	163	US-09-731-872-459	Sequence 459, App
2	463	100.0	163	US-09-876-997-459	Sequence 459, App
3	457	98.7	163	US-09-978-360A-668	Sequence 668, App
4	305	65.9	72	US-10-425-115-237341	Sequence 237341, App
5	134.5	29.0	281	US-10-424-599-237821	Sequence 237821, App
6	124	26.8	671	US-10-437-963-109378	Sequence 109378, App
7	123	26.8	359	US-10-425-115-231292	Sequence 231292, App
8	119.5	25.8	269	US-10-437-963-156973	Sequence 156973, App
9	117.5	25.4	242	US-10-425-115-254668	Sequence 254668, App
10	116.5	25.2	245	US-10-424-599-229595	Sequence 229595, App
11	116.5	25.2	263	US-10-437-963-152741	Sequence 152741, App
12	116	25.1	261	US-10-437-963-184232	Sequence 184232, App
13	116	25.1	265	US-10-425-115-201476	Sequence 201476, App

RESULT 2  
US-09-976-997-459  
; Sequence 459, Application US/09876997  
; Publication No. US20030152921A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bouquellet, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.USA.CIP  
; CURRENT APPLICATION NUMBER: US/09/876,997  
; PRIOR FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 09/731,872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 459  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-876-997-459

Query Match 100.0%; Score 463; DB 10; Length 163;  
Best Local Similarity 100.0%; Pred. No. 4e-43;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLGLALGRLGGSGARHVVFGS 60  
Db 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLGLALGRLGGSGARHVVFGS 60  
  
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91

RESULT 3  
US-09-978-360A-668  
; Sequence 668, Application US/09978360A  
; Publication No. US20040110939A1  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bouquellet, Lydie  
; APPLICANT: Jobert, Severin  
; APPLICANT: Clusel, Catherine  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
; FILE REFERENCE: 56.USA.CIP  
; CURRENT APPLICATION NUMBER: US/09/978,360A  
; PRIOR FILING DATE: 2004-10-15  
; PRIOR APPLICATION NUMBER: US 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: US 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: US 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: US 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: US 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: US 60/099,273  
; PRIOR FILING DATE: -09-04  
; PRIOR APPLICATION NUMBER: US 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: US 09/215,435  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: PCT/1998/02122  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: US 09/247,155

; PRIOR FILING DATE: 1999-02-09  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 810  
; SOFTWARE: Patent.pm  
; SEQ ID NO 668  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SIGNAL  
; LOCATION: -52...-1  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (36, 47, 54, 65, 70, 103, 107, 108)  
; OTHER INFORMATION: unknown  
US-09-978-360A-668  
  
Query Match 98.7%; Score 457; DB 11; Length 163;  
Best Local Similarity 98.9%; Pred. No. 1.9e-42;  
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLGLALGRLGGSGARHVVFGS 60  
Db 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLGLALGRLGGSGARHVVFGS 60  
  
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91  
  
RESULT 4  
US-10-425-115-237341  
; Sequence 237341, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 237341  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_148039C.1.pep  
US-10-425-115-237341  
  
Query Match 65.9%; Score 305; DB 17; Length 72;  
Best Local Similarity 86.1%; Pred. No. 4.7e-26;  
Matches 62; Conservative 2; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLGLALGRLGGSGARHVVFGS 60  
Db 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVDRGSKIRNLGLALGRLGGSGARHVVFGS 60  
  
QY 61 GRAAGKAVSCAE 72  
Db 61 GRAAGKAVSCAE 72  
  
RESULT 5  
US-10-424-599-237821  
; Sequence 237821, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K





Db 54 GRAINKTVMIAELIKRRIVGLHQT 78

## RESULT 9

US-10-425-115-254668  
; Sequence 254668, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 254668  
; LENGTH: 242  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(242)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: MRT4577\_163838C.1.pap  
US-10-425-115-254668

Query Match 25.4%; Score 117.5; DB 17; Length 242;  
Best Local Similarity 31.0%; Pred. No. 0.00011;  
Matches 27; Conservative 14; Mismatches 39; Indels 7; Gaps 1;  
QY 1 MEHYRKAGSVELPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSGARHVVFSGS 60  
Db 1 MDYQR-----VEKPRNETPIRENEIRITALGRNRYIGVGSMLLEENGHDIEITIKAM 53  
QY 61 GRAAGKAVSCAEIVKRVPGHLQTKL 87  
Db 54 GRAINKTVMIAELIKRRVAGLHQTSI 80

## RESULT 10

US-10-424-599-229595  
; Sequence 229595, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 229595  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_49349C.1.pap  
US-10-424-599-229595

Query Match 25.2%; Score 116.5; DB 15; Length 245;  
Best Local Similarity 31.0%; Pred. No. 0.00015;  
Matches 27; Conservative 16; Mismatches 37; Indels 7; Gaps 1;  
QY 1 MEHYRKAGSVELPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSGARHVVFSGS 60  
Db 1 MDYQR-----VEKPKAESPINENEIRVTSQGRTRNYITVATLLQKSGSSEIVLKAM 53

QY 61 GRAAGKAVSCAEIVKRVPGHLQTKL 87  
Db 54 GRAINKTVMIAELIKRRIVGLHQTQI 80

## RESULT 11

US-10-437-963-152741  
; Sequence 152741, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 152741  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_52762C.1.pap  
US-10-437-963-152741

Query Match 25.2%; Score 116.5; DB 16; Length 263;  
Best Local Similarity 29.4%; Pred. No. 0.00016;  
Matches 25; Conservative 15; Mismatches 38; Indels 7; Gaps 1;  
QY 1 MEHYRKAGSVELPAPSPMPQLPDTLEMRVDRGSKIRNLLGLALGRLEGSGARHVVFSGS 60  
Db 1 MDYQR-----VEKPRNETPISENEIRITAQGRMRNYISYGMSSLLENHGHDEIVIKAM 53  
QY 61 GRAAGKAVSCAEIVKRVPGHLQTL 85  
Db 54 GRAINKTVMIAELIKRRIGGLHQT 78

## RESULT 12

US-10-437-963-184232  
; Sequence 184232, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 184232  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(261)  
; OTHER INFORMATION: unsure at all Xaa locations

OTHER INFORMATION: Clone ID: LIB3356-013-D9\_FLI.pcp  
US-10-425-114-60641

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:47 ; Search time 23.1356 Seconds  
(without alignments)  
260.851 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_1\_91  
Perfect score: 463  
Sequence: 1 MEHYRKAGSVLPAPSPMPQ.....EIVKRRVPLGLTKLRLQ 91

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2.6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2.6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2.6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2.6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2.6/prodata/1/iaa/PTUS COMB.pep.\*  
6: /cgn2.6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	463	100.0	128	4	US-09-621-976-4077
2	395	85.3	78	4	US-09-513-999C-7680
3	151.5	32.7	118	4	US-09-270-767-39818
4	151.5	32.7	118	4	US-09-270-767-55035
5	75.5	16.3	261	4	US-09-489-039A-12408
6	73.5	15.9	271	3	US-08-968-563-33
7	73.5	15.9	271	3	US-08-969-683A-33
8	73.5	15.9	271	3	US-09-297-928-14
9	72.5	15.7	290	4	US-09-252-991A-123012
10	71	15.3	566	4	US-09-252-991A-18531
11	70	15.1	284	4	US-09-252-991A-17438
12	69	14.9	183	4	US-09-800-729-101
13	69	14.9	482	3	US-09-134-001C-4309
14	69	14.9	639	3	US-09-347-801-17
15	69	14.9	639	3	US-09-854-731-17
16	68.5	14.8	499	4	US-09-270-767-31754
17	68.5	14.8	499	4	US-09-270-767-46971
18	68	14.7	482	4	US-09-252-991A-23113
19	68	14.7	517	4	US-09-252-991A-24083
20	67.5	14.6	250	3	US-08-968-563-17
21	67.5	14.6	250	3	US-08-969-683A-17
22	67.5	14.6	250	3	US-09-297-928-13
23	67.5	14.6	250	4	US-09-641-652-56
24	67.5	14.6	250	4	US-09-538-092-240
25	67	14.5	70	4	US-09-489-847-375
26	67	14.5	402	4	US-09-252-991A-22463
27	1186	14.5	1186	2	US-08-861-464-8

28	67	14.5	1186	2	US-08-396-001-8	Sequence 8, Appli
29	67	14.5	1186	3	US-09-323-433A-8	Sequence 8, Appli
30	67	14.5	1186	4	US-09-826-752-8	Sequence 8, Appli
31	67	14.5	2595	3	US-09-036-987A-2	Sequence 2, Appli
32	67	14.5	2595	3	US-09-370-700-2	Sequence 2, Appli
33	67	14.5	2595	4	US-09-603-207-2	Sequence 2, Appli
34	66.5	14.4	496	3	US-09-134-001C-3001	Sequence 3001, Ap
35	66.5	14.4	555	4	US-09-252-991A-28848	Sequence 28848, A
36	66.5	14.4	580	4	US-09-252-991A-27245	Sequence 27245, A
37	65.5	14.1	240	4	US-09-252-991A-31100	Sequence 31100, A
38	65	14.0	877	2	US-08-916-917-2	Sequence 2, Appli
39	65	14.0	877	2	US-08-972-631-2	Sequence 2, Appli
40	65	14.0	877	2	US-08-972-629-2	Sequence 2, Appli
41	65	14.0	877	2	US-08-972-630-2	Sequence 2, Appli
42	65	14.0	877	2	US-08-672-211-2	Sequence 2, Appli
43	65	14.0	877	3	US-09-225-170-2	Sequence 2, Appli
44	64.5	13.9	244	1	US-08-696-827-1	Sequence 1, Appli
45	64.5	13.9	294	4	US-09-252-991A-29716	Sequence 29716, A

ALIGNMENTS

RESULT 1  
US-09-621-976-4077  
; Sequence 4077, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 4077  
; LENGTH: 128  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -52...-1  
; NAME/KEY: UNSURE  
; LOCATION: 72  
; OTHER INFORMATION: Xaa = Asp,Glu,Gly,Val  
US-09-621-976-4077

Query Match 100.0%; Score 463; DB 4; Length 128;  
Best Local Similarity 100.0%; Pred. No. 3e-50;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEHYRKAGSVLPAPSPMPQLPPTLEMRVROGSKIRNLLGLALGLEGGSARHVVFGS 60  
Db 1 MEHYRKAGSVLPAPSPMPQLPPTLEMRVROGSKIRNLLGLALGLEGGSARHVVFGS 60  
QY 61 GRAAGKAVSCAIVKRRVPLGLTKLRLQ 91  
Db 61 GRAAGKAVSCAIVKRRVPLGLTKLRLQ 91

RESULT 2  
US-09-513-999C-7680  
; Sequence 7680, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG

```
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7680
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-7680

Query Match      85.3%; Score 395; DB 4; Length 78;
Best Local Similarity 100.0%; Pred. No. 5.1e-42;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGSARHVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGSARHVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRV 78
Db 61 GRAAGKAVSCAEIVKRRV 78

RESULT 3
US-09-270-767-39818
; Sequence 39818, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39818
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39818

Query Match      32.7%; Score 151.5; DB 4; Length 118;
Best Local Similarity 41.8%; Pred. No. 2.8e-11;
Matches 33; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 11 ELPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGSARHVVFGSGRAAGKAVSC 70
Db 1 DLPFEDCMPKSKQDFLMHVKGTGKVSNIIEFAQALNKGHCRCVVMSSGGGVGKTISC 60
QY 71 AEIVKRRVGLHQLTKLRF 89
Db 61 AEVLKRSHPLXQVTRMAY 78

RESULT 4
US-09-270-767-55035
; Sequence 55035, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55035
; LENGTH: 118
; TYPE: PRT
```

```
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-55035

Query Match      32.7%; Score 151.5; DB 4; Length 118;
Best Local Similarity 41.8%; Pred. No. 2.8e-11;
Matches 33; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 11 ELPAPSPMPQLPDTLEMRVRDGSKIRNLLGLALGRLEGSARHVVFGSGRAAGKAVSC 70
Db 1 DLPFEDCMPKSKQDFLMHVKGTGKVSNIIEFAQALNKGHCRCVVMSSGGGVGKTISC 60
QY 71 AEIVKRRVGLHQLTKLRF 89
Db 61 AEVLKRSHPLXQVTRMAY 78

RESULT 5
US-09-489-039A-12408
; Sequence 12408, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12408
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12408

Query Match      16.3%; Score 75.5; DB 4; Length 261;
Best Local Similarity 27.7%; Pred. No. 0.26;
Matches 23; Conservative 18; Mismatches 31; Indels 11; Gaps 3;

QY 11 ELPAPSPMPQLPDTL-----EMRVRDGSKIRNLLGLALGRLEGSARHVVFGSGRAAG 65
Db 178 DLPKIVNPFQSDVLTLLAATLTWKEKATRPITMSNGK-SGVSRL-----VTGRULFG 231
QY 66 KAVSCAEIVKRRVPGHQLTKLR 88
Db 232 SAMTFTGVQSSAPQIAIAKLR 254

RESULT 6
US-08-968-563-33
; Sequence 33, Application US/08968563
; Patent No. 6013494
; GENERAL INFORMATION:
; APPLICANT: CHARLES E. NAKAMURA
; APPLICANT: ANTHONY A. GATENBY
; APPLICANT: AMY (KUANG-HUA) HSU
; APPLICANT: RICHARD D. LA REAU
; APPLICANT: SHARON L. HAYNIE
; APPLICANT: MARIA DIAZ-TORRES
; APPLICANT: DONALD E. TRIMBUR
; APPLICANT: GREGORY M. WHITED
; APPLICANT: VASANTHA NAGARAJAN
; APPLICANT: MARK S. PAYNE
; APPLICANT: STEPHEN K. PICATAGGIO
; APPLICANT: RAMESCH V. NAIR
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; PRODUCTION OF 1,3-PROPANEDIOL
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
```

STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
ADDRESSEE: GENENCOR INTERNATIONAL, INC.  
STREET: 4 CAMBRIDGE PLACE  
STREET: 1870 SOUTH WINTON ROAD  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 50 INCH DISKETTE  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95  
SOFTWARE: MICROSOFT WORD VERSION 7.0A  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968,563  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/030,601  
FILING DATE: NOVEMBER 13, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: FLOYD, LINDA AXAMETHY  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: CR-9982  
TELEPHONE: 302-892-8112  
TELEFAX: 302-773-0164  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE: GPPI  
ORGANISM: GPPI  
US-08-968-563-33

Query Match 15.9%; Score 73.5; DB 3; Length 271;  
Best Local Similarity 28.3%; Pred. No. 0.48;  
Matches 26; Conservative 13; Mismatches 32; Indels 21; Gaps 3;

QY 2 EHYRKAGSVELPAPSPMPOLPPDTLEMRVDRGSKIRNLLGLALGRLEGG SARHVF--SG 59  
Db 156 EYFITANDVKQKHPPEYL-----KGRNGLGFPINEQDPSKSKVVFDAP 202

QY 60 SGRAGKAVSCAEIVKRRVPGHLQLTKLRFQ 91  
Db 203 AGIAGKAAGC-----KIVGIATTFDLDLFLK 228

RESULT 7  
US-08-969-683A-33  
Sequence 33, Application US/08969683A  
Patent No. 6136576  
GENERAL INFORMATION:  
APPLICANT: GENENCOR INTERNATIONAL, INC.  
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT  
TITLE OF INVENTION: PRODUCTION OF 1,3 PROPANEDIOL  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 4 Cambridge Place  
STREET: 1870 South Winton road  
CITY: Rochester  
STATE: NY  
COUNTRY: U.S.A.  
ZIP: 14618  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,683A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/20873  
FILING DATE: 13-NOV-1997  
APPLICATION NUMBER: 60/030,601  
FILING DATE: 13-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Glaister, Debra  
REGISTRATION NUMBER: 33,888  
REFERENCE/DOCKET NUMBER: GC 369-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-864-7620  
TELEFAX: 650-845-6504  
TELEX:  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE: GPPI  
ORGANISM: GPPI  
US-08-969-683A-33

Query Match 15.9%; Score 73.5; DB 3; Length 271;  
Best Local Similarity 28.3%; Pred. No. 0.48;  
Matches 26; Conservative 13; Mismatches 32; Indels 21; Gaps 3;

QY 2 EHYRKAGSVELPAPSPMPOLPPDTLEMRVDRGSKIRNLLGLALGRLEGG SARHVF--SG 59  
Db 156 EYFITANDVKQKHPPEYL-----KGRNGLGFPINEQDPSKSKVVFDAP 202

QY 60 SGRAGKAVSCAEIVKRRVPGHLQLTKLRFQ 91  
Db 203 AGIAGKAAGC-----KIVGIATTFDLDLFLK 228

RESULT 8  
US-09-297-928-14  
Sequence 14, Application US/09297928  
Patent No. 6358716  
GENERAL INFORMATION:  
APPLICANT: BULTHUIS, BEN A.  
GATENBY, ANTHONY A.  
HAYNIE, SHARON L.  
HSU, AMY K.  
LAREAU, RICHARD D.  
TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF  
GLYCEROL BY RECOMBINANT  
ORGANISMS  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENENCOR INTERNATIONAL, INC.  
STREET: 4 CAMBRIDGE PLACE  
STREET: 1870 SOUTH WINTON ROAD  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 14618  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3 5 INCH  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MICROSOFT WINDOWS 95  
SOFTWARE: MICROSOFT WORD VERSION 7.0A  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/297,928

; FILING DATE: 11-May-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/03602  
; FILING DATE: NOVEMBER 13, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FLOYD, LINDA AXAMETHY  
; REGISTRATION NUMBER: 33,692  
; REFERENCE/DOCKET NUMBER: CR-9981-P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302-892-8112  
; TELEFAX: 302-773-0164  
; TELE: 6717325  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-297-928-14

Query Match 15.9%; Score 73.5; DB 3; Length 271;  
Best Local Similarity 28.3%; Pred. No. 0.48;  
Matches 26; Conservative 13; Mismatches 32; Indels 21; Gaps 3;  
QY 2 EHYRAGSVLPAPSPMPQLPPTDLEMRVDRGSKIRNLLGLALGRLEGSSARHVPF--SG 59  
Db 156 EYFITANDVKQKHPPEYL-----KGRNGLGFTINEQDPSSKVVVFEDAP 202  
QY 60 SGRAGKAVSCAEIVKRRVPGHQLTKLRFQ 91  
Db 203 AGIAGKAAGC-----KIVGIATTDFDLFLK 228

RESULT 9  
US-09-252-991A-23012  
; Sequence 23012, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23012  
; LENGTH: 290  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23012

Query Match 15.7%; Score 72.5; DB 4; Length 290;  
Best Local Similarity 34.8%; Pred. No. 0.7;  
Matches 23; Conservative 4; Mismatches 30; Indels 9; Gaps 2;  
QY 7 AGSVLPAPSPMPQLPPTDLEMRVDRGSKIRNLLGLALGRLEGSSARHVPFSGSRA 63  
Db 229 ADSPALFVQOPEHQHROPDRAGPACGAGPVVRPVG-----QPGEARHFSFLGGGTG 282  
QY 64 AGKAVS 69  
Db 283 AGKALS 288

RESULT 10  
US-09-252-991A-18531

; Sequence 18531, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18531  
; LENGTH: 566  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18531

Query Match 15.3%; Score 71; DB 4; Length 566;  
Best Local Similarity 28.7%; Pred. No. 2.6;  
Matches 25; Conservative 8; Mismatches 18; Indels 36; Gaps 5;  
QY 5 RKAGSVLPAP-----SPMPQL---PPTDLEMRVDRGSKIRNLLGLALGRLEGSSARH 54  
Db 87 QERGAVHYPRPLPAIPGDPHPRAGSWPDRLERLRQ----- 124  
QY 55 VVFGSG-GRAGKAVSCAEIVKRRVPG 80  
Db 125 --FPGAQGTATGRRRSQRIV-RQAPG 148

RESULT 11  
US-09-252-991A-17438  
; Sequence 17438, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17438  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17438

Query Match 15.1%; Score 70; DB 4; Length 284;  
Best Local Similarity 29.0%; Pred. No. 1.4;  
Matches 18; Conservative 10; Mismatches 32; Indels 2; Gaps 1;  
QY 12 LPAPSPMPQLPPTDLEMRVDRGSKIRNLLGLALGRLEGSSAR--HVVFSGSRAAGKAVS 69  
Db 26 LPSMDKRTQIPDPSFAARLKQAMAMRNLKQETLAEAGVQSNTIHKLTSGKAQSTRKLTJE 85  
QY 70 CA 71  
Db 86 IA 87

RESULT 12  
US-09-800-729-101  
; Sequence 101, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:



```
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (86)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (146)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-800-729-101

Query Match      14.9%; Score 69; DB 4; Length 183;
Best Local Similarity 32.8%; Pred. No. 1.1;
Matches 20; Conservative 5; Mismatches 16; Indels 20; Gaps 2;

QY 14 APSP---MPQLPPTLEMRVDSKIR-----RNLGLALGRLEGSGAR 53
DB 87 APSVMSMPAVEPDPVGRARPLRGLIVEIFYRCAGSCPRGARTQGLALARLQSGRX 146
QY 54 H 54
DB 147 H 147

RESULT 13
US-09-134-001C-4309
; Sequence 4309, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4309
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4309

Query Match      14.9%; Score 69; DB 3; Length 482;
Best Local Similarity 32.7%; Pred. No. 3.8;
Matches 17; Conservative 9; Mismatches 22; Indels 4; Gaps 1;

QY 21 LPPTLEMRVDSKIR----NLLGLALGRLEGSGARHVVFSGSGRAAGKAV 68
DB 7 IPPSKGEMNNNDTGDKRLRSTLSLIGLVIGSMICGGAFNIQSDMGHAGGLAI 58

RESULT 14
US-09-347-801-17
; Sequence 17, Application US/09347801
; Patent No. 6262345
; GENERAL INFORMATION:
```

```
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BB-1171
; CURRENT APPLICATION NUMBER: US/09/347,801
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,438
; EARLIER FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Zea mays
US-09-347-801-17

Query Match      14.9%; Score 69; DB 3; Length 639;
Best Local Similarity 25.6%; Pred. No. 5.4;
Matches 22; Conservative 15; Mismatches 33; Indels 16; Gaps 2;

QY 5 RKAGSVELPAPSPMPOLPPTLEMT-----RVRDGSKIRNLLGLALGRLEGSSA 52
DB 108 KEGAAANASPSRPRPQVKRVSSAGLLGSVLRRTENLKDKYSLGRRLGGQFGTT 167
QY 53 RHVVFSGSGRAAGKAVSCAEIVKRRV 78
DB 168 HLCV-----BRATGKELACKSILKRKL 189

RESULT 15
US-09-854-731-17
; Sequence 17, Application US/09854731
; Patent No. 6794561
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BB-1171
; CURRENT APPLICATION NUMBER: US/09/854,731
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/092,438
; PRIOR FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Zea mays
US-09-854-731-17

Query Match      14.9%; Score 69; DB 4; Length 639;
Best Local Similarity 25.6%; Pred. No. 5.4;
Matches 22; Conservative 15; Mismatches 33; Indels 16; Gaps 2;

QY 5 RKAGSVELPAPSPMPOLPPTLEMT-----RVRDGSKIRNLLGLALGRLEGSSA 52
DB 108 KEGAAANASPSRPRPQVKRVSSAGLLGSVLRRTENLKDKYSLGRRLGGQFGTT 167
QY 53 RHVVFSGSGRAAGKAVSCAEIVKRRV 78
DB 168 HLCV-----BRATGKELACKSILKRKL 189

Search completed: January 26, 2005, 15:08:52
Job time : 24.1356 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:09:58 ; Search time 39 Seconds  
(without alignments)  
224.506 Million cell updates/sec

Title: US-10-057-813-14 COPY 1 91

Accession: **U05406**  
 Length: **463**  
 Perfect score: **100**  
 Sequence: **1 MEHYRKAGSGVELPAPSPMPQ.....EIVRRVPGLHQLTKLRFLQ 91**

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 32090

```
Minimum DB seq length: 0
Maximum DB seq length: 91
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database :      PIR_79:*
1:  pir1:
2:  pir2:
3:  pir3:
4:  pir4:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	65.5	14.1	89	2	C69494	conserved hypothet	
2	60	13.0	72	2	I38497	anion exchanger 3	
3	56	12.1	89	2	C69383	conserved hypothet	
4	54.5	11.8	83	2	I73095	MHC class II beta	
5	51.5	11.1	79	2	C84802	hypothetical prote	
6	51.5	11.1	90	2	G75509	hypothetical prote	
7	51	11.0	87	2	T28356	ORF MSV195 ALI mot	
8	50.5	10.9	85	2	G84527	hypothetical prote	
9	50	10.8	88	2	E75144	hypothetical prote	
10	49	10.6	86	2	I83200	syntaxin 3E - mous	
11	49	10.6	86	2	H82855	hypothetical prote	
12	49	10.6	87	2	E64326	hypothetical prote	
13	48.5	10.5	83	2	AE2239	hypothetical prote	
14	48	10.4	80	2	AP2125	hypothetical prote	
15	48	10.4	84	2	S48815	hypothetical prote	
16	47.5	10.3	64	2	D96493	hypothetical prote	
17	47	10.2	73	2	AB1897	hypothetical prote	
18	47	10.2	75	2	G55891	hypothetical prote	
19	47	10.2	83	2	E65027	hypothetical prote	
20	47	10.2	86	2	I83199	syntaxin 3D - mous	
21	47	10.2	86	2	S08493	hypothetical prote	
22	47	10.2	88	1	WMAD9	early E1B 9K prote	
23	47	10.2	91	2	AB0804	conserved hypothet	
24	46	9.9	68	2	H72804	gp42 protein - Myc	
25	46	9.9	74	2	B40513	hypothetical prote	
26	46	9.9	80	2	E30010	hypothetical ORF-7	
27	46	9.9	87	2	D82136	cell division topo	
28	46	9.9	91	2	AB3204	conserved hypothet	
29	45.5	9.8	61	2	T34721	probable protoporp	

## ALIGNMENTS

## RESULT 1

conserved hypothetical protein AF1956 - Archaeoglobus fulgidus  
C/Species: Archaeoglobus fulgidus  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Aug-2004  
C/Accession: C69494

C;Accession: C03474  
R;Klien, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F  
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A;Reference number: A69250; MUID:98049343; PMID:9389475  
A;Accession: C69494  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-89 <KLE>  
C;Cross-references: UNIPROT:O28323; GB:AE000968; GB:AE000782; NID:Q2699291; PIDN:AA8892  
C;Superfamily: Archaeal histone-like protein, Alba type

Query Match	14.1%	Score 65.5;	DB 2;	Length 89;
Best Local Similarity	41.2%;	Pred. NO. 6.3;		
Matches	21; Conservative	8; Mismatches	19; Indels	3; Gaps 3;
Qy	33	GSK-IRNLGLALGRLEGSSARHVPVSGSGRAGKAVSCAEIVKKR-VPGIL	81	
	:	: : : : : :	:	:
Dd	9	GNEPVNMYVLATLTOLNEG-ADEWJLKARGRAISRAVDVAEIVRNEMPCGV	58	
	:	: : : : : :	:	:

## RESULT 2

anion exchanger 3 cardiac isoform - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
 C:Accession: I38497  
 R:Yannoukakos, D.; Stuart-Tilley, A.; Fernandez, H.; Fey, P.; Duyk, G.; Alper, S.  
 Circ. Res. 75, 603-614, 1994  
 A:Title: Molecular cloning, expression, and chromosomal localization of two isofo  
 A:Reference number: I38496; MUID:95008042; PMID:7923606  
 A:Accession: I38497  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-72 <RES>  
 A:Cross-references: UNIPROT:Q12835; EMBL:U05597; NID:9476223; PIDN:AAA50749.1; PI  
 C:Superfamily: band 3 anion transport protein  
 C:Keywords: cardiac muscle; heart

Query Match 13.0%; Score 60; DB 2; Length 72;  
Best Local Similarity 43.6%; Pred. No. 19;  
Matches 17; Conservative 1; Mismatches 15; Indels 6; Gaps 1;

QY 13 PAPSPMPQLPDDTLEMRVD-----GSKIRNLGLALG 45  
 ||||| : : : : :  
 Db 31 PAPSPRSLAPDELMFVLPFDGDLWESIRGLQGPWAG 69  
 ||||| : : : : :  
 RESULT 3  
 C69383  
 conserved hypothetical protein AF1067 - Archaeoglobus fulgidus  
 C/Species: Archaeoglobus fulgidus  
 C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Aug-2004  
 C/Accession: C69383  
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea  
 A/Reference number: A69250; MUID:98049343; PMID:9389475  
 A/Accession: C69383  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-89 <KLE>  
 A/Cross-references: UNIPROT:Q29195; GB:AE001029; GB:AE000782; NID:g2689352; PIDN:AAB9017  
 C/Suprafamily: Archaeal histone-like protein, Alba type  
 Query Match 12.1%; Score 56; DB 2; Length 89;  
 Best Local Similarity 37.0%; Pred. No. 63;  
 Matches 17; Conservative 7; Mismatches 20; Indels 2; Gaps 2;  
 QY 33 GSK-INNLGLALGRLEGGSRHVFSGGAAAGKAVSCAEIVKRR 77  
 ||| : : : : :  
 Db 9 GNRKPVNMYVLAVLTQFNSG-ATEVSIKARGRAISRAVDVAEIVKRR 53  
 ||| : : : : :  
 RESULT 4  
 I73095  
 MHC class II beta chain - slender loris (fragment)  
 C/Species: Loris tardigradus (slender loris)  
 C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 C/Accession: I73095  
 R/Pignero, F.; O'Huigin, C.; Tichy, H.; Klein, J.  
 J. Immunol. 152, 4455-4465, 1994  
 A/Title: The origin of the primate Mhc-DRB genes and allelic lineages as deduced from th  
 A/Reference number: I56227; MUID:94209665; PMID:8157963  
 A/Accession: I73095  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-83 <RES>  
 A/Cross-references: UNIPROT:Q30547; UNIPROT:Q30543; EMBL:Z27126; NID:g415838; PIDN:CAA81  
 A/Note: DRB\*W2005  
 A/Accession: I73099  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-83 <RE2>  
 A/Cross-references: EMBL:Z27161; NID:g415844; PID:g415845  
 A/Note: DRB\*W2006  
 Query Match 11.8%; Score 54.5; DB 2; Length 83;  
 Best Local Similarity 26.0%; Pred. No. 84;  
 Matches 20; Conservative 12; Mismatches 34; Indels 11; Gaps 2;  
 QY 7 AGSVLPAPSPMPQLPDDTLEMRVDRGSKIRNLGLALGRLEGGSRHVFSGGAAAGK 66  
 ||| : : : : :  
 Db 18 AGAV-----PGETPLPGGERPLQRGG-----GVSGDGDGAGAAVRVLEQPEGRPGGG 66  
 ||| : : : : :  
 QY 67 AVSCAEIVKRRVPLHQ 83  
 ||| : : : : :  
 Db 67 AGRGGHVLTQPLDLRR 83  
 ||| : : : : :  
 RESULT 5  
 C84802

hypothetical protein At2g38210 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C/Accession: C84802  
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A/Reference number: A84420; MUID:20083487; PMID:10617197  
 A/Accession: C84802  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-79 <STO>  
 A/Cross-references: UNIPROT:O80446; GB:AE002093; NID:g3335369; PIDN:AAC27170.1; GSPDB:G  
 C/Genetics:  
 A/Gene: At2g38210  
 A/Map position: 2  
 Query Match 11.1%; Score 51.5; DB 2; Length 79;  
 Best Local Similarity 40.0%; Pred. No. 1.7e+02;  
 Matches 16; Conservative 5; Mismatches 18; Indels 1; Gaps 1;  
 QY 40 LGLALGRLEGGSRHVFSGGAAAGKAVSCAEIVKRRVP 79  
 ||||| : : : : :  
 Db 28 VGLA-QMLRGVIMDVNREQAIAEEAGACAVMALERVP 66  
 ||||| : : : : :  
 RESULT 6  
 G75509  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C/Species: Deinococcus radiodurans  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C/Accession: G75509  
 R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A/Reference number: A75250; MUID:20036896; PMID:10567266  
 A/Accession: G75509  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-90 <WHI>  
 A/Cross-references: UNIPROT:Q9RWZ8; GB:AE001910; GB:AE000513; NID:g6458198; PIDN:AAF1005  
 A/Experimental source: strain R1  
 C/Genetics:  
 A/Gene: DR0517  
 A/Map position: 1  
 Query Match 11.1%; Score 51.5; DB 2; Length 90;  
 Best Local Similarity 35.4%; Pred. No. 1.9e+02;  
 Matches 17; Conservative 4; Mismatches 14; Indels 13; Gaps 2;  
 QY 5 RKAGSVEL-----PAPSPMPQLPDDTLEMRVDRGSKIRNLGLALGR 47  
 ||| : : : : :  
 Db 27 RKTASTKTAASSNTAPAPQHP-----SGSRAPLILGLALAVL 66  
 ||| : : : : :  
 RESULT 7  
 T28356  
 ORF MSV195 ALI motif gene family protein - Melanoplus sanguinipes entomopoxvirus  
 C/Species: Melanoplus sanguinipes entomopoxvirus  
 C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
 C/Accession: T28356  
 R/Afonso, C.L.; Tullman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.  
 J. Virol. 73, 533-552, 1999  
 A/Title: The genome of Melanoplus sanguinipes entomopoxvirus.  
 A/Reference number: Z20484; MUID:99102612; PMID:9847359  
 A/Accession: T28356  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-87 <AFO>



; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1059-1073, 1996  
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
A;Reference number: A64300; MUID:96337999; PMID:8688087  
A;Accession: E64326  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-87 <BUL>  
A;Cross-references: UNIPROT:Q57665; GB:U67477; GB:L77117; NID:gl590949; PIDN:AA898197.1;  
C;Genetics:  
A;Map position: FOR203759-204022  
C;Superfamily: Archaeal histone-like protein, Alba type

Query Match	10.6%	Score 49;	DB 2;	Length 87;
Best Local Similarity	39.1%	Pred. NO. 3.4e+02;		
Matches	9;	Conservative	5;	Mismatches 9;
				Indels 0;
Qy	55	VVFGSGRAAGKAVSCAEIVKRR	77	
	:	:	:	:
Db	29	VIIKARGKAINKAVDVAEMIRN	51	
	:	:	:	:

RESULT 13  
AE2239  
hypothetical protein asr3468 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-JUL-2004  
C:Accession: AE2239  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Kakezaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA res. 8, 205-213, 2001  
A:Title: Complete genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: ABI807; MUID:21595285; PMID:11759840  
A:Accession: AE2239  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-83 <RUR>  
A:Cross-references: UNIPROT:Q8YRH9; GB:BA000019; PIDN:BAB75167.1; PID:gl7132601; GSPDB:C  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: asr3468

```

Query Match      10.5%   Score 48.5;   DB 2;   Length 83;
Best Local Similarity 28.3%;   Pred. No. 3.6e+02;
Matches 15;   Conservative 5;   Mismatches 14;   Indels 19;   Gaps 1;

29  RVRCGSKTRNLGLALGRLEGSGARHVFGSGRAAGKAVSCAEIVKRRVPEGL 81
||||| : : : : :
19  RVRDEE-----VIIISQAGTFPIARVIAEQKLPRIPGEL 52

```

RESULT 14  
AF2125  
Proteomorphological protein asl2557 [imported] - Nostoc sp. (strain PCC 7120)  
PCC 7120  
Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
Accession: AF2125  
C:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
NAKAZAKI, N., 2005-213, 2001  
DNA Res. 8, 205-211  
Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
Reference number: AB1807; MUID:21595285; PMID:11759840  
Accession: AF2125  
Status: preliminary  
Molecule type: DNA  
Residues: 1-80 <KUR>  
Cross-References: UNIPROT:Q8YU04; GB:BA000019; PIDN:BAB74256.1; PID:gl7131649; GSFD8:G  
Experimental source: strain PCC 7120  
Genetics:

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:08:58 ; Search time 193 Seconds  
(without alignments)

271.291 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_1\_91

Perfect score: 463

Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPLGLHQLKRLFLQ 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 239185

Minimum DB seq length: 0

Maximum DB seq length: 91

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	65.5	14.1	89	1	ALB2 ARCFU	O28323 archaeoglob
2	60	13.0	72	2	Q12835	Q12835 homo sapien
3	58	12.5	71	2	Q9J9W4	Q9J9W4 tadpole ede
4	58	12.5	71	2	Q9J9W8	Q9J9W8 bufo marinu
5	58	12.5	71	2	Q9J9X0	Q9J9X0 bufo marinu
6	58	12.5	71	2	Q9J9X2	Q9J9X2 leptodactyl
7	58	12.5	71	2	Q9J9X6	Q9J9X6 frog virus
8	58	12.5	71	2	Q9J9Y2	Q9J9Y2 bufo marinu
9	58	12.5	71	2	Q9J9Y4	Q9J9Y4 bufo marinu
10	58	12.5	71	2	Q9J9Z6	Q9J9Z6 bufo marinu
11	58	12.5	71	2	Q9J9Z8	Q9J9Z8 rana tempor
12	58	12.5	71	2	Q9J9A0	Q9J9A0 rana tempor
13	57.5	12.4	74	2	Q9SFC3	Q9SFC3 arabisopsis
14	56	12.1	77	2	Q92MC1	Q92MC1 rhizobium m
15	56	12.1	89	1	ALB1 ARCFU	O29195 archaeoglob
16	55.5	12.0	64	2	O8S3X2	O8S3X2 oryza sativ
17	55.5	12.0	86	2	O6ZSD5	O6ZSD5 oryza sativ
18	55.5	12.0	86	2	BAD10112	BAD10112 oryza sat
19	55.5	12.0	91	1	ALBA METTH	O27527 methanobact
20	55	11.9	91	2	O6Z896	O6Z896 oryza sativ
21	55	11.9	91	2	BAD07703	BAD07703 oryza sat
22	55	11.9	91	2	BAD07903	BAD07903 oryza sat
23	54.5	11.8	83	2	Q30543	Q30543 loris tardi
24	54.5	11.8	83	2	Q30547	Q30547 loris tardi
25	54	11.7	52	1	PH68 HUMAN	Q9uht9 homo sapien
26	54	11.7	52	2	AAH69658	AAH69658 homo sapi
27	54	11.7	52	2	AAH69757	AAH69757 homo sapi
28	54	11.7	52	2	AAH69775	AAH69775 homo sapi
29	53.5	11.6	71	2	Q9J9Y8	Q9J9Y8 bufo bufo u
30	53.5	11.6	71	2	Q9J9Z0	Q9J9Z0 bufo bufo u
31	53.5	11.6	71	2	Q9J9Z2	Q9J9Z2 bufo bufo u

32 53.5 11.6 73 2 Q6Z8S3  
33 53.5 11.6 73 2 BAD10039  
34 53 11.4 90 2 Q99PM8  
35 53 11.4 91 2 Q96T24  
36 52.5 11.3 59 2 Q61JF3  
37 52 11.2 88 2 Q6FVB9  
38 52 11.2 91 2 Q8GH93  
39 51.5 11.1 79 1 PXL4 ARATH  
40 51.5 11.1 90 2 Q8RWZ8  
41 51 11.0 71 2 Q9J9W6  
42 51 11.0 71 2 Q9J9X4  
43 51 11.0 71 2 Q9J9Y0  
44 51 11.0 71 2 Q9J9Y6  
45 51 11.0 71 2 Q9J9Z4

Q6Z8S3 oryza sativ  
BAD10039 oryza sat  
Q99PM8 mus musculus  
Q96T24 neurospora  
Q61JF3 drosophila  
Q6FVB9 candida gla  
Q8GH93 dictyosteli  
Q80446 arabidopsis  
Q9RWZ8 deinococcus  
Q9J9W6 sheatfish i  
Q9J9X4 guppyfish i  
Q9J9Y0 doctor fish  
Q9J9Y6 catfish iri  
Q9J9Z4 bohle irido

#### ALIGNMENTS

##### RESULT 1

ALB2 ARCFU

ID ALB2 ARCFU STANDARD; PRT; 89 AA.

AC O28323;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 01-OCT-2004 (Rel. 45, Last annotation update)

DE DNA/RNA-binding protein Alba 2 (Afl).

GN Name=alba2; OrderedLocusNames=AF1956;

OS Archaeoglobus fulgidus.

OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;

OC Archaeoglobaceae; Archaeoglobus.

NCBI\_TaxID=2234;

[1] SEQUENCE FROM N.A.

RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;

RX MEDLINE=98049433; PubMed=9389475; DOI=10.1038/37052;

RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,

RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,

RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,

RA Kirkness E.P., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,

RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,

RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,

RA Utterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P.,

RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,

RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,

Woese C.R., Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-

reducing archaeon Archaeoglobus fulgidus.";

Nature 390:364-370(1997).

[2]

RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS), AND MUTAGENESIS OF ASN-10;

LYS-11; LEU-18 AND PHE-54.

RX MEDLINE=22726758; PubMed=12730210; DOI=10.1074/jbc.M303666200;

RA Zhao K., Chai X., Marmorstein R.;

RT "Structure of a Sir2 substate. Alba, reveals a mechanism for

deacetylation-induced enhancement of DNA binding.";

RL J. Biol. Chem. 278:26071-26077(2003).

CC -!- FUNCTION: Binds double-stranded DNA tightly but without sequence

specificity. It is distributed uniformly and abundantly on the

chromosome, suggesting a role in chromatin architecture. However,

it does not significantly compact DNA. Binds rRNA and mRNA in

vivo. May play a role in maintaining the structural and functional

stability of RNA, and, perhaps, ribosomes (By similarity).

CC -!- SUBUNIT: Homotetramer.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -!- PTM: Acetylated. Deacetylation by the Sir2-homolog deacetylase may

regulate its activity (By similarity).

CC -!- SIMILARITY: Belongs to the archaeal histone-like Alba family.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its





```
DE Hypothetical protein (Fragment).
OS Bufo marinus Venezuelan iridovirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus;
OC Bufo iridoviruses.
OX NCBI_TaxID=105554;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157674; AAF64585.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMVRVDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQPLPPSPKLRFRTRDNOEV 63

RESULT 6
Q9J9X2
ID Q9J9X2 PRELIMINARY; PRT; 71 AA.
AC Q9J9X2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Lepidactylus Venezuelan iridovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Ranavirus.
OX NCBI_TaxID=100219;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157672; AAF64583.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMVRVDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQPLPPSPKLRFRTRDNOEV 63

RESULT 7
Q9J9X6
ID Q9J9X6 PRELIMINARY; PRT; 71 AA.
AC Q9J9X6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Frog virus 3 (FV3).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Ranavirus.
OX NCBI_TaxID=104993;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157668; AAF64579.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMVRVDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQPLPPSPKLRFRTRDNOEV 63

RESULT 8
Q9J9Y2
ID Q9J9Y2 PRELIMINARY; PRT; 71 AA.
AC Q9J9Y2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Bufo marinus Venezuelan iridovirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus;
OX NCBI_TaxID=105555;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157662; AAF64573.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMVRVDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQPLPPSPKLRFRTRDNOEV 63

RESULT 9
Q9J9Y4
ID Q9J9Y4 PRELIMINARY; PRT; 71 AA.
AC Q9J9Y4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Bufo marinus Venezuelan iridovirus 6.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus;
OX NCBI_TaxID=105556;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
```

```

DR EMBL; AF157660; AAF64571.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match
Best Local Similarity 12.5%; Score 58; DB 2; Length 71;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQLPPSPKLRFRTRDQEV 63

RESULT 10
Q9J926 PRELIMINARY; PRT; 71 AA.
ID Q9J926 AC
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Bufo marinus Venezuela iridovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OC Bufo iridoviruses.
OX NCBI_TaxID=105552;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157648; AAF64559.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match
Best Local Similarity 12.5%; Score 58; DB 2; Length 71;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQLPPSPKLRFRTRDQEV 63

RESULT 11
Q9J928 PRELIMINARY; PRT; 71 AA.
ID Q9J928 AC
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rana temporaria United Kingdom iridovirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=105558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157646; AAF64557.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match
Best Local Similarity 12.5%; Score 58; DB 2; Length 71;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQLPPSPKLRFRTRDQEV 63

RESULT 12
Q9JA00 PRELIMINARY; PRT; 71 AA.
ID Q9JA00 AC
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rana temporaria United Kingdom iridovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=105557;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157644; AAF64555.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match
Best Local Similarity 12.5%; Score 58; DB 2; Length 71;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQLPPSPKLRFRTRDQEV 63

RESULT 13
Q9SFC3 PRELIMINARY; PRT; 74 AA.
ID Q9SFC3 AC
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE F17A17.25 protein (AT3907910/F17A17_25) (Hypothetical protein).
GN Name=F17A17.25;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eutrosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
RA Banh J., Bowler L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M., Satou M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

```

RN [3]  
 RX MEDLINE=22088475; PubMed=12093376;  
 RA Haas B.J., Volfovsky N., Town C.D.,  
 RA Feldmann K.A., Flavell R., White O., Salzberg S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation.";  
 RL genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayaishizaki Y.,  
 RA Iehida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC013483; AAF21201.1; -;  
 DR EMBL; AY060549; AAL31180.1; -;  
 DR EMBL; AY086394; AAM64461.1; -;  
 DR EMBL; AF412117; AAL06569.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 74 AA; 7368 MW; 44410B581F4D51ED CRC64;  
 Query Match 12.4%; Score 57.5; DB 2; Length 74;  
 Best Local Similarity 32.0%; Pred. No. 3e+02;  
 Matches 16; Conservative 7; Mismatches 12; Indels 15; Gaps 2;  
 QY 41 GLALGRLEGCSARHVFSGSGRAGKAVCAEIVKRRVPGHQLTKRLFL 90  
 Db 13 GVAVGGALGG-----AVGAVGYTYEARVKVPLGLH---KVRFI 47  
 RESULT 14  
 Q92MC1 ID Q92MC1 PRELIMINARY; PRT; 77 AA.  
 AC Q92MC1;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE HYPOTHETICAL TRANSMEMBRANE PROTEIN.  
 GN OrderedLocusNames=R02706; ORFNames=SMC00666;  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21368234; PubMed=11474104;  
 RA Galibert F., Finan T.M., Long S.R., Puehler A., Boistard P., Bothe G.,  
 RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Chain P.,  
 RA Boutry M., Bowser L., Brehmer J., Cadieu E., Capela D., Chain P.,  
 RA Cowie A., Davis R.W., Dreano S., Federpiel N.A., Fisher R.F.,  
 RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,  
 RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,  
 RA Kahn M.L., Kalman S., Keating D.H., Kies E., Komp C., Lelaure V.,  
 RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,  
 RA Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,  
 RA Vorhoeiter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;  
 RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";  
 RL Science 293:668-672(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RN

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kies E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL; AL591791; CAC47285.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 KW Complete proteome; Hypothetical protein; Transmembrane.  
 SQ SEQUENCE 77 AA; 8618 MW; 2CEA7D9103F1EAEB CRC64;  
 Query Match 12.1%; Score 56; DB 2; Length 77;  
 Best Local Similarity 46.2%; Pred. No. 4.5e+02;  
 Matches 18; Conservative 4; Mismatches 13; Indels 4; Gaps 2;  
 QY 15 PSMPLQPPDTLEMR---VRDGSKIRNLGLALGRLEGG 50  
 Db 32 PQPAERNPKQTLPRGQGLRFLSPRLNWLGVAL-MLVGG 69  
 RESULT 15  
 ALBI\_ARCFU ID ALBI\_ARCFU STANDARD; PRT; 89 AA.  
 AC O29135;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE DNA/RNA-binding protein Alba 1 (Af2).  
 GN Name=albi1; OrderedLocusNames=AF1067;  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OX NCBI\_TaxID=2234;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyrpides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,  
 RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,  
 RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,  
 RA Uterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P.,  
 RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,  
 RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,  
 RA Woese C.R., Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997).  
 CC -1- FUNCTION: Binds double-stranded DNA tightly but without sequence  
 CC specificity. It is distributed uniformly and abundantly on the  
 CC chromosome, suggesting a role in chromatin architecture. However,  
 CC it does not significantly compact DNA. Binds rRNA and mRNA in  
 CC vivo. May play a role in maintaining the structural and functional  
 CC stability of RNA, and, perhaps, ribosomes (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBUNIT LOCATION: Cytoplasmic (Probable).  
 CC -1- PTM: Acetylated. Deacetylation by the SIR2-homolog deacetylase may  
 CC regulate its activity (By similarity).  
 CC -1- SIMILARITY: Belongs to the archaeal histone-like Alba family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

Search completed: January 26, 2005, 15:34:39  
Job time : 194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:07:38 ; Search time 156 Seconds  
(without alignments)  
209.259 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_1\_91  
Perfect score: 463  
Sequence: 1 MEHYRAGSVELPAPSPMPQ.....EIVKRRVPLHQLTKLRFQ 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 segs, 358729299 residues

Total number of hits satisfying chosen parameters: 1094936

Minimum DB seq length: 0  
Maximum DB seq length: 91

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : 'A\_Geneseq\_23Sep04:.'  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	402	86.8	79	2	AAY11950 Human 5'
2	395	85.3	78	3	AAG03599 Human sec
3	119	25.7	74	3	AAG36978 Arabidops
4	67	14.5	70	3	AAY91690 Human sec
5	67	14.5	70	8	ADL1771 Novel hum
6	62.5	13.5	63	8	ADN99364 Novel hum
7	62	13.4	84	8	ADG22762 Cyanophag
8	61.5	13.3	86	3	AAG34486 Arabidops
9	57.5	12.4	71	3	AAG58753 Arabidops
10	57.5	12.4	74	3	AAG15942 Arabidops
11	57.5	12.4	85	4	ABG00727 Novel hum
12	55	11.9	75	5	ABP10404 Human ORF
13	54	11.7	83	4	AAG64347 Cu/Zn SOD
14	53.5	11.6	67	4	ABB43040 Peptide #
15	53.5	11.6	67	4	AAM36870 Peptide #
16	53.5	11.6	67	4	AAM76763 Human bon
17	53.5	11.6	67	4	AAM63945 Human bra
18	53.5	11.6	67	4	ABG58446 Human liv
19	53.5	11.6	88	4	ABG18684 Novel hum
20	53	11.4	81	4	AAM86389 Human imm
21	53	11.4	83	8	ADO20465 Human PRO
22	52.5	11.3	64	4	AAM84673 Human imm
23	52.5	11.3	64	4	AAM57537 Propionib
24	52.5	11.3	64	6	ABM54056 Propionib
25	52.5	11.3	76	3	AAG53244 Arabidops

26	52.5	11.3	77	4	AAG73945 Human col
27	52.5	11.3	77	4	AAU64906 Propionib
28	52.5	11.3	77	6	ABM61425 Propionib
29	52	11.2	71	8	ABO58355 Human gen
30	52	11.2	80	4	AAU53651 Propionib
31	52	11.2	80	6	ABM50170 Propionib
32	52	11.2	81	4	AAM95345 Human tes
33	52	11.2	81	4	ABB96039 Human tes
34	52	11.2	87	4	AAM92262 Human dig
35	51.5	11.1	61	4	AAU53022 Propionib
36	51.5	11.1	61	6	ABM49541 Propionib
37	51.5	11.1	65	3	AAG31725 Arabidops
38	51.5	11.1	79	3	AAG31723 Arabidops
39	51	11.0	72	6	ABP80069 N. gonorr
40	51	11.0	75	4	AAM13691 Peptide #
41	51	11.0	75	4	ABB32622 Peptide #
42	51	11.0	75	4	AAM26091 Peptide #
43	51	11.0	75	4	ABB27471 Human pep
44	51	11.0	75	4	ABB18120 Protein #
45	51	11.0	75	4	AAM65829 Human bon

ALIGNMENTS

RESULT 1  
AAY11950  
ID AAY11950 standard; protein; 79 AA.  
XX AC AAY11950;  
XX  
XX 18-JUN-1999 (first entry)  
XX  
XX Human 5' EST secreted protein SEQ ID No: 550.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;  
KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;  
KW upstream regulatory sequence; cytokine activity; cell proliferation;  
KW differentiation; haematopoiesis regulation; tissue growth regulation;  
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
KW thrombolytic; anti-inflammatory; tumour inhibition.  
XX  
XX Homo sapiens.  
XX  
XX WO9906550-A2.  
XX  
XX 11-FEB-1999.  
XX  
XX 31-JUL-1998; 98WO-IB001232.  
XX  
XX 01-AUG-1997; 97US-00905144.  
XX  
XX (GEST ) GENSET.  
XX  
XX Dumas Milne Edwards J, Duclert A, Lacroix B;  
XX  
XX WPI: 1999-153780/13.  
XX  
XX N-PSDB; AAX40672.

XX New isolated prostate-derived nucleic acids - used to develop products  
XX which may have cytokine, immune regulatory, haematopoiesis regulating,  
XX anti-inflammatory or tumour inhibition activity.

XX Claim 34; Page 648; 675pp; English.

XX AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for  
XX human secreted proteins expressed in prostate, and encode the proteins  
XX given in AAY11716 to AAY11993 respectively. The proteins given represent  
XX the signal peptide and an N-terminal fragment of a secreted protein. The  
XX nucleic acid sequences can be used for producing secreted human gene  
XX products. They can also be used to develop products for diagnosis and  
XX therapy. The proteins obtained may have cytokine activity, cell  
XX proliferation and differentiation activity, haematopoiesis regulating

CC activity, tissue growth regulating activity, reproductive hormone  
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and  
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
CC activity, tumour inhibition activity or other activities. The products  
CC can be used in forensic, gene therapy and chromosome mapping procedures.  
CC The sequences can also be used for obtaining corresponding promoter  
CC sequences. The nucleic acids encoding the signal peptides can be used for  
CC directing extracellular secretion of a polypeptide or the insertion of a  
CC polypeptide into a membrane, or importing a polypeptide into a cell  
XX  
SQ Sequence 79 AA;

Query Match 86.8%; Score 402; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 5.9e-40;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60  
Db |||||  
1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60  
QY 61 GRAAGKAVSCAEIVKRRVP 79  
Db |||||  
61 GRAAGKAVSCAEIVKRRVP 79

RESULT 2  
AAG03599  
ID AAG03599 standard; protein; 78 AA.

XX AAG03599;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 7680.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI: 2000-500381/45.

XX N-PSDB; AAC03605.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 13; SEQ ID NO 7680; 71pp + Sequence Listing; English.

XX The present sequence is a polypeptide encoded by one of a large number of  
CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
CC prepared from total human RNAs or polyA+ RNAs derived from 30 different  
CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
CC region (UTR) of the mRNA because they are often obtained from oligo-dT  
CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
CC sequences derived from the 5' ends of mRNAs and even in those cases where  
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely  
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can  
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
CC are also used in diagnostic, forensic, gene therapy and chromosome  
CC mapping procedures. They are used to obtain upstream regulatory sequences

CC and to design expression and secretion vectors  
XX  
SQ Sequence 78 AA;

Query Match 85.3%; Score 395; DB 3; Length 78;  
Best Local Similarity 100.0%; Pred. No. 4e-39;  
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60  
Db |||||  
1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60

QY 61 GRAAGKAVSCAEIVKRRV 78

Db |||||  
61 GRAAGKAVSCAEIVKRRV 78

RESULT 3

AAG36978

ID AAG36978 standard; protein; 74 AA.

XX AAG36978;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 45395.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126284P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132485P.

XX 07-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 18-MAY-1999; 99US-0134370P.

XX 19-MAY-1999; 99US-0134768P.

XX 20-MAY-1999; 99US-0134941P.

XX 21-MAY-1999; 99US-0135124P.

XX 24-MAY-1999; 99US-0135353P.

XX 25-MAY-1999; 99US-0135629P.

XX 27-MAY-1999; 99US-0136021P.

XX 99US-0136392P.

PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 06-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-01377502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 17-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139452P.	PR 13-AUG-1999;	99US-0148565P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149222P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149222P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149229P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-01400353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-01400354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 20-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159637P.
PR 21-JUL-1999;	99US-0144814P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0145086P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160741P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160814P.
PR 23-JUL-1999;	99US-0145193P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160981P.
PR 26-JUL-1999;	99US-0145276P.	PR 22-OCT-1999;	99US-0160989P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145919P.	PR 25-OCT-1999;	99US-0161406P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161359P.
PR 02-AUG-1999;	99US-0146338P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146388P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161920P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161992P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0161993P.
		PR 29-OCT-1999;	99US-0162142P.





DR WPI; 2004-180094/17.  
 XX New human secreted nucleic acid, useful for diagnosing and treating  
 PT neurodegenerative, inflammatory, hyperproliferative, metabolic,  
 PT reproductive, cardiovascular, respiratory or immunological disorders or  
 PT diseases.  
 XX Disclosure; SEQ ID NO 375; 234pp; English.  
 XX  
 CC The invention describes an isolated human nucleic acid molecule (1)  
 CC comprising a polynucleotide having a nucleotide sequence at least 95%  
 CC identical to: a sequence polynucleotide fragment of SEQ ID NO: X or of  
 CC the cDNA sequence included in ATCC Deposit No: Z, which is hybridisable  
 CC to SEQ ID NO: X; or a sequence encoding a polypeptide fragment, domain or  
 CC epitope of SEQ ID NO: Y or a polypeptide sequence encoded by the cDNA  
 CC sequence included in ATCC Deposit No: Z, which is hybridisable to SEQ ID  
 CC NO: X, having a biological activity. The nucleic acids and polypeptides,  
 CC pharmaceutical formulations and kits are useful in diagnosing and  
 CC treating neurodegenerative diseases states, behavioral disorders,  
 CC inflammatory conditions, hyperproliferative disorders (e.g. Alzheimer's  
 CC disease, Parkinson's disease or Huntington's disease), metabolic  
 CC disorders (e.g. Tay-Sachs disease or Leish-Nyhan syndrome), reproductive  
 CC disorders, immunological disorders (e.g. arthritis, asthma or AIDS),  
 CC endocrine and immune disorders (e.g. Hodgkin's lymphoma), haematopoietic  
 CC or muscular disorders (e.g. leukaemia), autoimmune disorders, allergy,  
 CC cancer, cardiovascular, respiratory or pulmonary disorders, disorders or  
 CC conditions affecting connective tissue, skin disorders, CNS disorders,  
 CC congenital disorders, infectious disorders and gastrointestinal  
 CC disorders. This is the amino acid sequence of a novel human secreted  
 CC protein fragment of the invention. Note: this sequence does not appear in  
 CC the printed specification but is available in electronic format from the  
 CC US patent office at [ftp.seqdata.uspto.gov/seqdata.html?docID=20040034196](http://ftp.seqdata.uspto.gov/seqdata.html?docID=20040034196).  
 XX  
 SQ Sequence 70 AA;  
 Query Match 14.5%; Score 67; DB 8; Length 70;  
 Best Local Similarity 41.0%; Pred. No. 4.5;  
 Matches 16; Conservative 2; Mismatches 9; Indels 12; Gaps 1;  
 QY 13 PAPSPQLPPDTLEMRVDRGSKIRNLGLALGRLEGG 51  
 DB 10 PLPSPALAP-----AHSLLGLLGRMGSS 36  
 RESULT 6  
 ADN99364  
 ID ADN99364 standard; protein; 63 AA.  
 AC ADN99364;  
 XX 29-JUL-2004 (first entry)  
 DT Novel human protein sequence #180.  
 DE anti-inflammatory; dermatological; neuroprotective; immunomodulator;  
 KW antibacterial; virucide; antipsoriatic; cytostatic; gene therapy;  
 KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;  
 KW early aging; hormonal imbalance; ischemic heart disease;  
 KW ulcerative colitis.  
 OS Homo sapiens.  
 XX WO2004038003-A2.  
 PN 06-MAY-2004.  
 XX 24-OCT-2003; 2003WO-US033947.  
 XX 25-OCT-2002; 2002US-0421061P.  
 PR 25-OCT-2002; 2002US-0421080P.  
 PR 25-OCT-2002; 2002US-0421522P.  
 PR 25-OCT-2002; 2002US-0421614P.  
 PR 30-OCT-2002; 2002US-0422177P.

PR 30-OCT-2002; 2002US-0422178P.  
 PR 15-NOV-2002; 2002US-0426355P.  
 PR 15-NOV-2002; 2002US-0426384P.  
 PR 15-NOV-2002; 2002US-0426394P.  
 PR 15-NOV-2002; 2002US-0426430P.  
 PR 15-NOV-2002; 2002US-0426916P.  
 PR 27-NOV-2002; 2002US-0429224P.  
 PR 27-NOV-2002; 2002US-0429275P.  
 PR 27-NOV-2002; 2002US-0429302P.  
 PR 27-NOV-2002; 2002US-0429326P.  
 PR 27-NOV-2002; 2002US-0429511P.  
 PR 04-DEC-2002; 2002US-0430845P.  
 PR 04-DEC-2002; 2002US-0430851P.  
 PR 04-DEC-2002; 2002US-0430657P.  
 PR 04-DEC-2002; 2002US-0430663P.  
 PR 04-DEC-2002; 2002US-0430668P.  
 PR 04-DEC-2002; 2002US-0430884P.  
 PR 05-DEC-2002; 2002US-0430937P.  
 PR 05-DEC-2002; 2002US-0430965P.  
 PR 05-DEC-2002; 2002US-0431458P.  
 PR 13-DEC-2002; 2002US-0433251P.  
 PR 13-DEC-2002; 2002US-0433500P.  
 PR 13-DEC-2002; 2002US-0433316P.  
 PR 13-DEC-2002; 2002US-0433318P.  
 PR 23-DEC-2002; 2002US-0436236P.  
 PR 03-JAN-2003; 2003US-0437914P.  
 PR 17-JAN-2003; 2003US-0440820P.  
 PR 17-JAN-2003; 2003US-0440821P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 03-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476621P.  
 PR 09-JUN-2003; 2003US-0476632P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485217P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 08-JUL-2003; 2003US-0485359P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;  
 PI Wong JGP, Wu G, Zhang H, Zeng C;  
 XX WPI; 2004-365511/34.  
 DR N-PSDB; ADN98580.  
 XX New nucleic acid molecules, useful in preparing a composition for  
 PT treating or preventing e.g. inflammatory, CNS, bacterial or viral  
 PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or  
 PT ulcerative colitis.  
 XX

PS Claim 14; SEQ ID NO 964; 532pp; English.

XX The invention relates to a nucleic acid molecule comprising a

CC polynucleotide sequence or its complement that encodes a polypeptide. The

CC nucleic acid is useful in preparing a composition for treating or

CC preventing inflammatory, CNS, immune, bacterial or viral disorder,

CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic

CC heart disease or ulcerative colitis. This sequence corresponds to a

CC protein of the invention.

XX

SQ Sequence 63 AA;

Query Match 13.5%; Score 62.5; DB 8; Length 63;

Best Local Similarity 40.7%; Pred. No. 14;

Matches 22; Conservative 5; Mismatches 10; Indels 17; Gaps 4;

Qy 7 AGSV-ELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEG-----GSARH 54

Db 13 AGSIRHRPPASAPPPR-PPQL-----SRDLTGTCRLKGRPNGRARH 55

RESULT 7

ADG22762

ID ADG22762 standard; protein; 84 AA.

XX

AC ADG22762;

XX

DT 26-FEB-2004 (first entry)

XX

DE Cyanophages S-2L encoded protein #507.

XX

KW genome; cyanophages; 2; 6-diaminopurine; chemotherapy; AIDS.

XX

OS Cyanophages S-2L.

XX

PN FR2839079-A1.

XX

PD 31-OCT-2003.

XX

PF 30-APR-2002; 2002FR-00005424.

XX

PR 30-APR-2002; 2002FR-00005424.

XX

PA (INSP ) INST PASTEUR.

PA (CNRS ) CNRS CENT NAT RECH SCI.

PA (GENO-) GENOSCOPE CENT NAT SEQUENCE GRP INTERE.

XX

PI Marliere P, Kaminski PA, Galisson F, Bouzon M, Pochet S;

PI Weissenbach J, Saurin W, Robert C, Vico V;

XX

DR WPI; 2004-045746/05.

DR N-PSDB; ADG22255.

XX

PT New genomic sequence for cyanophages S-2L, useful for identifying genes

PT for synthesis of 2,6-diaminopurine bases or polynucleotides containing

PT them.

XX

PS Claim 6; SEQ ID NO 508; 423pp; French.

XX

CC The invention relates to the entire genome of cyanophages S-2L, and to the

CC protein encoded by it. Genes isolated from the genome of S-2L are useful

CC for preparing enzymes for synthesis of D-bases (D = 2,6-diaminopurine),

CC particularly D, dTMP and dDTP, or polynucleotides containing these bases,

CC polymerases involved in metabolism of D-bases and deoxynucleotide

CC analogs, for chemotherapy of AIDS. The genes, and encoded polypeptides,

CC can be used for detection and/or identification of S-2L, and for

CC identifying agents that modulate synthesis of D-bases or polynucleotides

CC containing them, and fusions of S-2L polypeptides with an antigen can be

CC used to raise specific antibodies, useful for detecting S-2L. This

CC sequence corresponds to one of the proteins encoded by the cyanophages S-

CC 2L genome.

XX

SQ Sequence 84 AA;

Query Match 13.4%; Score 62; DB 8; Length 84;

Best Local Similarity 29.3%; Pred. No. 22;

Matches 22; Conservative 11; Mismatches 26; Indels 16; Gaps 3;

Qy 7 AGSVELPAP-----SPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFG 59

Db 12 AGOMVLPAFPGGLDRGPLEDEEQGVAMGVLPFRVVPVAGVGARLEQG-----FAG 65

Qy 60 SGRAGKAVSCAIV 74

Db 66 GGH---RAVTRREAV 77

RESULT 8

AAG34486

ID AAG34486 standard; protein; 86 AA.

XX

AC AAG34486;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 41968.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 18-MAY-1999; 99US-0134370P.

PR 19-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139494P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 13.3%; Score 61.5; DB 3; Length 86;  
Best Local Similarity 36.5%; Pred. No. 26;

```
Matches 19; Conservative 4; Mismatches 18; Indels 11; Gaps 2;
QY 14 APSMPOLPPTTLEMRVDSKIRNLGLALGRLEGGSARHVVFGSGSRAAG 65
Db 45 APAVKPCPTDTLKLGVFGG-----LGGYGGGEGGG-----YGGSGGGGG 85

RESULT 9
AAG58753
ID AAG58753 standard; protein; 71 AA.
XX
AC AAG58753;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 75896.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132040P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 07-MAY-1999; 99US-0132486P.
PR 11-MAY-1999; 99US-0132487P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140821P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
```

PR	13-AUG-1999;	99US-0148565P.	RESULT 10	
PR	13-AUG-1999;	99US-0148684P.	AAG15942	
PR	16-AUG-1999;	99US-0149368P.	ID	AAG15942 standard; protein; 74 AA.
PR	17-AUG-1999;	99US-0149175P.	XX	
PR	18-AUG-1999;	99US-0149426P.	AC	AAG15942;
PR	20-AUG-1999;	99US-0149722P.	XX	
PR	20-AUG-1999;	99US-0149723P.	DT	17-OCT-2000 (first entry)
PR	20-AUG-1999;	99US-0149929P.	XX	
PR	23-AUG-1999;	99US-0149930P.	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 16390.
PR	25-AUG-1999;	99US-0150566P.	KW	Protein identification; signal transduction pathway; metabolic pathway;
PR	25-AUG-1999;	99US-0150884P.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	27-AUG-1999;	99US-0151065P.	KW	termination sequence.
PR	27-AUG-1999;	99US-0151066P.	XX	
PR	27-AUG-1999;	99US-0151080P.	OS	Arabidopsis thaliana.
PR	30-AUG-1999;	99US-0151303P.	XX	
PR	31-AUG-1999;	99US-0151438P.	PN	EP1033405-A2.
PR	01-SEP-1999;	99US-0151930P.	XX	
PR	07-SEP-1999;	99US-0152363P.	PD	06-SEP-2000.
PR	10-SEP-1999;	99US-0153070P.	XX	
PR	13-SEP-1999;	99US-0153758P.	PF	25-FEB-2000; 2000EP-00301439.
PR	15-SEP-1999;	99US-0154018P.	XX	
PR	16-SEP-1999;	99US-0154039P.	XX	25-FEB-1999; 99US-0121825P.
PR	20-SEP-1999;	99US-0154779P.	PR	05-MAR-1999; 99US-0123180P.
PR	22-SEP-1999;	99US-0155539P.	PR	09-MAR-1999; 99US-0123548P.
PR	23-SEP-1999;	99US-01555486P.	PR	23-MAR-1999; 99US-0125788P.
PR	24-SEP-1999;	99US-0155659P.	PR	25-MAR-1999; 99US-0126264P.
PR	28-SEP-1999;	99US-0156458P.	PR	29-MAR-1999; 99US-0126785P.
PR	29-SEP-1999;	99US-0156596P.	PR	01-APR-1999; 99US-0127462P.
PR	04-OCT-1999;	99US-0157117P.	PR	06-APR-1999; 99US-0128234P.
PR	05-OCT-1999;	99US-0157753P.	PR	08-APR-1999; 99US-0128714P.
PR	06-OCT-1999;	99US-0157865P.	PR	16-APR-1999; 99US-0129845P.
PR	07-OCT-1999;	99US-0158029P.	PR	19-APR-1999; 99US-0130077P.
PR	08-OCT-1999;	99US-0158232P.	PR	21-APR-1999; 99US-0130449P.
PR	12-OCT-1999;	99US-0158369P.	PR	23-APR-1999; 99US-0130510P.
PR	13-OCT-1999;	99US-0159293P.	PR	23-APR-1999; 99US-0130891P.
PR	13-OCT-1999;	99US-0159294P.	PR	28-APR-1999; 99US-0131449P.
PR	13-OCT-1999;	99US-0159295P.	PR	30-APR-1999; 99US-0132048P.
PR	14-OCT-1999;	99US-0159329P.	PR	30-APR-1999; 99US-0132407P.
PR	14-OCT-1999;	99US-0159330P.	PR	04-MAY-1999; 99US-0132484P.
PR	14-OCT-1999;	99US-0159331P.	PR	05-MAY-1999; 99US-0132485P.
PR	14-OCT-1999;	99US-0159637P.	PR	06-MAY-1999; 99US-0132486P.
PR	14-OCT-1999;	99US-0159638P.	PR	06-MAY-1999; 99US-0132487P.
PR	18-OCT-1999;	99US-0159584P.	PR	07-MAY-1999; 99US-0132863P.
PR	21-OCT-1999;	99US-0160741P.	PR	11-MAY-1999; 99US-0134256P.
PR	21-OCT-1999;	99US-0160767P.	PR	14-MAY-1999; 99US-0134218P.
PR	21-OCT-1999;	99US-0160768P.	PR	14-MAY-1999; 99US-0134219P.
PR	21-OCT-1999;	99US-0160770P.	PR	14-MAY-1999; 99US-0134221P.
PR	21-OCT-1999;	99US-0160814P.	PR	14-MAY-1999; 99US-0134370P.
PR	21-OCT-1999;	99US-0160815P.	PR	18-MAY-1999; 99US-0134768P.
PR	22-OCT-1999;	99US-0160980P.	PR	19-MAY-1999; 99US-0134941P.
PR	22-OCT-1999;	99US-0160981P.	PR	20-MAY-1999; 99US-0135124P.
PR	22-OCT-1999;	99US-0160989P.	PR	21-MAY-1999; 99US-0135353P.
PR	23-OCT-1999;	99US-0161404P.	PR	24-MAY-1999; 99US-0135629P.
PR	25-OCT-1999;	99US-0161405P.	PR	25-MAY-1999; 99US-0136021P.
PR	25-OCT-1999;	99US-0161406P.	PR	27-MAY-1999; 99US-0136392P.
PR	26-OCT-1999;	99US-0161359P.	PR	28-MAY-1999; 99US-0136782P.
PR	26-OCT-1999;	99US-0161360P.	PR	01-JUN-1999; 99US-0137222P.
PR	26-OCT-1999;	99US-0161361P.	PR	03-JUN-1999; 99US-0137528P.
PR	28-OCT-1999;	99US-0161920P.	PR	04-JUN-1999; 99US-0137502P.
PR	28-OCT-1999;	99US-0161922P.	PR	07-JUN-1999; 99US-0137724P.
PR	28-OCT-1999;	99US-0161993P.	PR	08-JUN-1999; 99US-0138094P.
PR	29-OCT-1999;	99US-0162142P.	PR	10-JUN-1999; 99US-0138540P.
PR			PR	10-JUN-1999; 99US-0138847P.
PR			PR	14-JUN-1999; 99US-0139119P.
PR			PR	16-JUN-1999; 99US-0139452P.
PR			PR	16-JUN-1999; 99US-0139453P.
PR			PR	17-JUN-1999; 99US-0139492P.
PR			PR	18-JUN-1999; 99US-0139454P.
PR			PR	18-JUN-1999; 99US-0139455P.
PR			PR	18-JUN-1999; 99US-0139456P.
PR			PR	18-JUN-1999; 99US-0139457P.
PR			PR	18-JUN-1999; 99US-0139458P.

Query Match 12.4%; Score 57.5; DB 3; Length 71;  
Best Local Similarity 32.4%; Pred. No. 62;  
Matches 16; Conservative 7; Mismatches 12; Indels 15; Gaps 2;  
QY 41 GLALGRLEGGSARRVFGSGRAAGKAVSCAEIVKRRVPGCLHQLTKLREL 90  
DB 13 GVAVGALGG-----AVGAVYGYEAIKRVKPGHL---KVRFI 47



```

XX DE Novel human diagnostic protein #718.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS64914.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 31086; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activities. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
XX CC patent did not appear in the printed specification. Note: The sequence data for this
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 85 AA;

Query Match 12.4%; Score 57.5; DB 4; Length 85;
Best Local Similarity 32.3%; Pred. No. 77;
Matches 21; Conservative 6; Mismatches 23; Indels 15; Gaps 3;

QY 5 RKAGSVELPA--PSPMPQLPDTLEMRVDSGKIRNLGLALGRLEGSGARHVFGSGR 62
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 33 RIAGPPVPVAGGPPPPRAP-----RAREGA-----CPAARRDRGPPVSALRPKRG 79
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 63 AAGKA 67
   |||||
Db 80 AAGKA 84
   |||||

RESULT 12
ABP10404
ID ABP10404 standard; protein; 75 AA.
XX AC ABP10404;

```

```

XX DT 24-JUN-2002 (first entry)
XX DE Human ORFX protein sequence SEQ ID NO:20790.
XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX KW hypertension; hypothyroidism; cholesterol ester storage disease;
XX KW immune deficiency; immune disorder; infectious disease;
XX KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX KW myasthenia gravis.
XX OS Homo sapiens.
XX PN WO200192523-A2.
XX PD 06-DEC-2001.
XX PF 29-MAY-2001; 2001WO-US010836.
XX PR 30-MAY-2000; 2000US-0206132P.
XX PR 29-AUG-2000; 2000US-0228716P.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach MD;
XX DR WPI; 2002-106308/14.
XX DR N-PSDB; ABN26156.
XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
XX PT preventing and treating cardiovascular disease, neurodegenerative,
XX PT hyperproliferative disorders and autoimmune disorders.
XX PS Disclosure; SEQ ID NO 20790; 1037pp; English.
XX CC The present invention describes substantially purified human proteins
XX CC (referred to as open reading frame, ORFX where X is 1-11491 (see Table 1
XX CC in the specification)). ABN15762 to ABN27252 encode the human ORFX
XX CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
XX CC treating or preventing a pathology associated with an ORFX-associated
XX CC disorder in humans, and in the manufacture of a medicament for treating a
XX CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
XX CC sequences can be used in gene therapy. ORFX sequences can be used in the
XX CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
XX CC osteoarthritis, neurodegenerative disorders, disorders related to organ
XX CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
XX CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
XX CC storage disease, various immune deficiencies and disorders, infectious
XX CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
XX CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
XX CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
XX CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
XX CC bone degenerative disorders, or periodontal disease, and for gut
XX CC protection or regeneration and treatment of lung or liver fibrosis,
XX CC reperfusion injury in various tissues and conditions resulting from
XX CC systemic cytokine damage. N.B. The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 75 AA;

Query Match 11.9%; Score 55; DB 5; Length 75;
Best Local Similarity 44.4%; Pred. No. 1.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 62 RAAGKAVSCAEIVKRVVPGHQLTKUR 88
   ||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 27 RREGKQTRCAHVYGLSLPGLNLTKER 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```





```

PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
PT
XX Claim 27; SEQ ID NO 37139; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs:
CC see AA131315-AA157346). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
XX Sequence 67 AA;
SQ
    Query Match          11.6%; Score 53.5; DB 4; Length 67;
    Best Local Similarity 30.8%; Pred. No. 1.7e+02;
    Matches 16; Conservative 8; Mismatches 11; Indels 17; Gaps 3;

QY      8 GSVELP-----ASPMQPLPPDTLEWRVDSKIR-----NLGLALGRLE 48
        : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      9 GLIDYPLGWAPSPKPMPELP-----MGDNTKSQAMSANFLGSLVTLTQQ 54

```

Search completed: January 26, 2005, 15:31:21  
Job time : 160 secs

***This Page Blank (uspto)***



```

RESULT 2
US-10-424-599-186121
; Sequence 186121, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 186121
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_13907C.1.pep
; US-10-424-599-186121

```

[illegible]

```

RESULT 3
US-10-351-334-375
; Sequence 375, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P203192
; CURRENT APPLICATION NUMBER: US/10/351,334
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 375
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-351-334-375

```

Query Match 14.5%; Score 67; DB 15; Length 70;  
Best Local Similarity 41.0%; Pred. No. 10;  
Matches 16; Conservative 2; Mismatches 9; Indels 12; Gaps 1;

Qy 13 PAPSPQLPPDTLEMRVDSKIRNLGLALGRLEGG 51  
| | | | |  
Db 10 PLPSAPAPALP-----AHSULGLLGRMSGSS 36  
: | | | | |

## RESULT 4

```

US-10-425-115-240946
; Sequence 240946, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 240946
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_151321C.1.pep
US-10-425-115-240946

```

```

Query Match      14.4%; Score 66.5; DB 17; Length 79;
Best Local Similarity 28.8%; Pred. No. 13;
Matches 21; Conservative 8; Mismatches 29; Indels 15; Gaps 1;

QY 10 VELPAPSPMPOLPPDTTLEMRVRDGSKIRNLLGLALGRLEGGSAHRVVFSGSGRAAGKAVS 69
Db 12 VPLPYLPSPSTGPWN-----ANALGTNFGGGAAGVFGGAGFQADRAVT 56

QY 70 CAEIVKRRVPGLH 82
Db 57 FLPLKRGTFIGIH 69

```

## RESULT 5

```

US-10-437-963-141539
; Sequence 141539, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 141539
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_42632C.1.pep
US-10-437-963-141539

```

Query Match 13.3%; Score 61.5; DB 16; Length 81;  
Best Local Similarity 38.3%; Pred. No. 48;  
Matches 23; Conservative 6; Mismatches 22; Indels 9; Caps 3;



Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 333500  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_67265C.1.pep  
US-10-425-115-333500

Query Match 12.4%; Score 57.5; DB 17; Length 87;  
Best Local Similarity 34.4%; Pred. No. 1.5e+02;  
Matches 22; Conservative 9; Mismatches 18; Indels 15; Gaps 5;  
Qy 2 EHVKAGSVEL---PAPSP---MPQLPP-DTLEMVRDGSKIRN-----LLGLALGRLEG 49  
Db 22 EHKRSSPHCELHTCPTAGNLMPLPHLPWPWSLALPQEGRGCTSSPVLIGLAVG---G 78  
Qy 50 GSAR 53  
Db 79 GEER 82

RESULT 11  
US-10-425-115-362864  
; Sequence 362864, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 362864  
; LENGTH: 61  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_94106C.1.pep  
US-10-425-115-362864

Query Match 12.3%; Score 57; DB 17; Length 61;  
Best Local Similarity 30.3%; Pred. No. 1.1e+02;  
Matches 20; Conservative 2; Mismatches 18; Indels 26; Gaps 2;  
Qy 14 APSMPQLPPDTLEMVRDGSKIRNLLGLALGRLEGGSARHVVFSGSRAAGKAVSCAEI 73  
Db 6 AGFPCTLGFST-----ALEVGPCRH-----GPASSRAVPCLGT 39  
Qy 74 VKRRVP 79  
Db 40 VGRAVP 45

RESULT 12  
US-10-437-963-147885  
; Sequence 147885, Application US/10437963  
; Publication No. US20040123343A1

; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 147885  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_48370C.1.pep  
US-10-437-963-147885

Query Match 12.3%; Score 57; DB 16; Length 86;  
Best Local Similarity 45.7%; Pred. No. 1.6e+02;  
Matches 16; Conservative 4; Mismatches 13; Indels 2; Gaps 1;  
Qy 45 GRL--EGGSARHVVFSGSRAAGKAVSCAEIVKRR 77  
Db 45 GRLGEGGQARNGAASGLGRDGYRMERGEVDGRQ 79

RESULT 13  
US-10-437-963-113150  
; Sequence 113150, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 113150  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(87)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_16966C.1.pep  
US-10-437-963-113150

Query Match 12.3%; Score 57; DB 16; Length 87;  
Best Local Similarity 32.2%; Pred. No. 1.7e+02;  
Matches 19; Conservative 3; Mismatches 31; Indels 6; Gaps 1;  
Qy 3 HYRKAGSVLPAPSPMPQLPPDTLEMVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 61  
Db 21 HCHCALVSSSPSPPPPTPLTSLRUSRPSPLVESLG-----EGRKAEIEQESAGG 73  
RESULT 14

US-10-424-599-190837  
; Sequence 190837, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 190837  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(90)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_14345C.1.pep  
US-10-424-599-190837

Query Match 12.1%; Score 56; DB 15; Length 90;  
Best Local Similarity 27.1%; Pred. No. 2.2e+02;  
Matches 19; Conservative 10; Mismatches 29; Indels 12; Gaps 4;

QY 15 PSMPPQI-----PP-----DTLEMRV-RDGSKIR-----NLLGLALGRLEGGSGARHVVFGSGR 62  
| | | | | : : : : : | | | | | : : : : :  
Db 5 PRPRTLTHMQPSPSYXQCYQWRIAKHGGRPKFHHINTRHCALLRXHGGPPLHIVYGWVGK 64

QY 63 AACKAVSCAE 72  
: | :  
Db 65 KSAKPFHCQ 74

RESULT 15  
US-10-437-963-107873  
; Sequence 107873, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Bing  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 107873  
; LENGTH: 91  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(91)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_12182C.1.pep  
US-10-437-963-107873

Query Match 12.1%; Score 56; DB 16; Length 91;  
Best Local Similarity 37.0%; Pred. No. 2.3e+02;  
Matches 17; Conservative 4; Mismatches 15; Indels 10; Gaps 2;

QY 26 LEMRVRDGSKIRNLLGLALGRLEGGSGARHVVFGSGRAAGKAVSCA 71  
: | | | | : : : : : | | | | | : : : : :  
Db 12 IEQRRRVGTH-----GVARGRVVGGAGH-----GPRATGAGECA 47

Search completed: January 26, 2005, 15:47:04  
Job time : 145 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:24:36 ; Search time 38 Seconds  
(without alignments)

158.814 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_1\_91

Perfect score: 463

Sequence: 1 MEHYRKAGSVLPAPSPMPQ.....EIVKRVPGHLHQLTKLRFQ 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 301706

Minimum DB seq length: 0

Maximum DB seq length: 91

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*

2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*

3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*

4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*

5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pcp.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	395	85.3	78	4	US-09-513-999C-7680
2	67	14.5	70	4	US-09-489-847-375
3	52	11.2	69	4	US-09-621-976-6159
4	50.5	10.9	81	4	US-09-371-671B-6
5	49.5	10.7	85	3	US-09-430-564-8
6	49.5	10.7	85	3	US-09-430-564-14
7	49	10.6	85	4	US-09-489-039A-13911
8	48.5	10.5	87	4	US-09-248-796A-16579
9	48	10.4	45	3	US-09-018-635-37
10	48	10.4	45	4	US-09-012-962-37
11	48	10.4	61	4	US-09-540-236-3294
12	48	10.4	71	4	US-09-270-767-58217
13	48	10.4	80	3	US-09-134-001C-2843
14	47	10.2	59	4	US-09-513-999C-6477
15	47	10.2	60	4	US-09-270-767-37413
16	47	10.2	60	4	US-09-270-767-52630
17	47	10.2	77	3	US-08-803-346-62
18	47	10.2	77	3	US-08-803-346-63
19	47	10.2	87	4	US-09-513-999C-6225
20	47	10.2	90	4	US-09-134-000C-3428
21	46.5	10.0	36	6	5451527-6
22	46.5	10.0	41	6	5451527-9
23	46.5	10.0	41	6	5451527-10
24	46.5	10.0	41	6	5451527-11
25	46.5	10.0	61	4	US-08-469-260A-507
26	46.5	10.0	61	4	US-08-488-446-507
27	46.5	10.0	61	4	US-08-467-344A-507

28	46.5	10.0	61	4	US-09-583-110-3458	Sequence 3458, Ap
29	46.5	10.0	61	4	US-08-424-550B-507	Sequence 507, App
30	46.5	10.0	64	4	US-09-621-976-6713	Sequence 6713, Ap
31	46.5	10.0	90	4	US-09-621-976-4538	Sequence 4538, Ap
32	46	9.9	53	4	US-09-257-179-88	Sequence 88, Appl
33	46	9.9	76	6	5202239-11	Patent No. 5202239
34	45.5	9.8	72	4	US-09-734-492A-6	Sequence 6, Appl
35	45.5	9.8	76	1	US-08-519-777-22	Sequence 22, Appl
36	45.5	9.8	76	1	US-08-742-035-22	Sequence 22, Appl
37	45.5	9.8	76	2	US-08-777-019-22	Sequence 22, Appl
38	45.5	9.8	76	2	US-08-777-143-22	Sequence 22, Appl
39	45.5	9.8	76	3	US-08-775-414-22	Sequence 22, Appl
40	45.5	9.8	76	3	US-08-931-858B-22	Sequence 22, Appl
41	45.5	9.8	76	3	US-08-981-739-22	Sequence 22, Appl
42	45.5	9.8	76	4	US-09-128-026-22	Sequence 22, Appl
43	45.5	9.8	76	4	US-09-220-616-22	Sequence 22, Appl
44	45.5	9.8	76	4	US-09-220-527-22	Sequence 22, Appl
45	45.5	9.8	76	4	US-09-220-407-22	Sequence 22, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-513-999C-7680  
; Sequence 7680, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCES: 59.US2.REG  
; PATENT NO. 6783961  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7680  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: Homo sapiens

##### Result

No.	Score	Query Match	Length	DB ID	Description
1	395	85.3	78	4	US-09-513-999C-7680
2	67	14.5	70	4	US-09-489-847-375
3	52	11.2	69	4	US-09-621-976-6159
4	50.5	10.9	81	4	US-09-371-671B-6
5	49.5	10.7	85	3	US-09-430-564-8
6	49.5	10.7	85	3	US-09-430-564-14
7	49	10.6	85	4	US-09-489-039A-13911
8	48.5	10.5	87	4	US-09-248-796A-16579
9	48	10.4	45	3	US-09-018-635-37
10	48	10.4	45	4	US-09-012-962-37
11	48	10.4	61	4	US-09-540-236-3294
12	48	10.4	71	4	US-09-270-767-58217
13	48	10.4	80	3	US-09-134-001C-2843
14	47	10.2	59	4	US-09-513-999C-6477
15	47	10.2	60	4	US-09-270-767-37413
16	47	10.2	60	4	US-09-270-767-52630
17	47	10.2	77	3	US-08-803-346-62
18	47	10.2	77	3	US-08-803-346-63
19	47	10.2	87	4	US-09-513-999C-6225
20	47	10.2	90	4	US-09-134-000C-3428
21	46.5	10.0	36	6	5451527-6
22	46.5	10.0	41	6	5451527-9
23	46.5	10.0	41	6	5451527-10
24	46.5	10.0	41	6	5451527-11
25	46.5	10.0	61	4	US-08-469-260A-507
26	46.5	10.0	61	4	US-08-488-446-507
27	46.5	10.0	61	4	US-08-467-344A-507

Query Match 85.3%; Score 395; DB 4; Length 78;  
Best Local Similarity 100.0%; Pred. No. 5.1e-42;  
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEHYRKAGSVLPAPSPMPOLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVFSGS 60  
Db 1 MEHYRKAGSVLPAPSPMPOLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVFSGS 60

Qy 61 GRAAGKAVSCAEIVKRRV 78  
Db 61 GRAAGKAVSCAEIVKRRV 78

##### RESULT 2

US-09-489-847-375  
; Sequence 375, Application US/09489847  
; Patent No. 6476195  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al  
; TITLE OF INVENTION: 98 Human Secreted Proteins  
; FILE REFERENCE: P2031P1  
; CURRENT APPLICATION NUMBER: US/09/489,847  
; CURRENT FILING DATE: 2000-01-24  
; EARLIER APPLICATION NUMBER: PCT/US99/17130  
; EARLIER FILING DATE: 1999-07-29  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30

; EARLIER APPLICATION NUMBER: 60/095,486  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER APPLICATION NUMBER: 60/096,319  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER APPLICATION NUMBER: 60/095,454  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER APPLICATION NUMBER: 60/095,455  
; EARLIER FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 376  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 375  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-489-847-375

Query Match 14.5%; Score 67; DB 4; Length 70;  
Best Local Similarity 41.0%; Pred. No. 0.54;  
Matches 16; Conservative 2; Mismatches 9; Indels 12; Gaps 1;  
QY 13 PAPSPMPQLPDDTLEMRVDSKIRNLLGLALGREGGS 51  
DB 10 PLPSPAPALP-----AHSLLGLLGRMSGSS 36

RESULT 3  
US-09-621-976-6159  
; Sequence 6159, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6159  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 29  
; OTHER INFORMATION: Xaa = Ala,Pro  
; NAME/KEY: UNSURE  
; LOCATION: 32  
; OTHER INFORMATION: Xaa = Cys,Gly,Arg,Ser  
US-09-621-976-6159

Query Match 11.2%; Score 52; DB 4; Length 69;  
Best Local Similarity 31.2%; Pred. No. 40;  
Matches 15; Conservative 5; Mismatches 22; Indels 6; Gaps 1;  
QY 3 HYRKAGSVLPAPSPMPQLPDDTLEMRVDSKIRNLLGLALGREGG 50  
DB 26 HHRXPCHXCSAPGPAP-----ARHATWGGQDVHSHYGRREGG 67

RESULT 4  
US-09-371-671B-6  
; Sequence 6, Application US/09371671B  
; Patent No. 6548743  
; GENERAL INFORMATION:  
; APPLICANT: Sheen, Jen  
; APPLICANT: Chiu, Wan-Ling  
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING A  
; FILE REFERENCE: 00786/370002  
; CURRENT APPLICATION NUMBER: US/09/371,671B  
; CURRENT FILING DATE: 1999-08-10

; PRIOR APPLICATION NUMBER: 60/155,934  
; PRIOR FILING DATE: 1999-01-14  
; PRIOR APPLICATION NUMBER: 60/095,938  
; PRIOR FILING DATE: 1998-08-10  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 81  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-371-671B-6

Query Match 10.9%; Score 50.5; DB 4; Length 81;  
Best Local Similarity 35.4%; Pred. No. 76;  
Matches 17; Conservative 7; Mismatches 21; Indels 3; Gaps 1;  
QY 44 LGRLEGGSRHVVFSSGSGRAAGKAVSCAIEVKK---RVPGLHQLTKLR 88  
DB 16 LDQCEKDKARVLVHCVMSGKSRSPAVVAVYLMKRGWRLAESHOWYKQR 63

RESULT 5  
US-09-430-564-8  
; Sequence 8, Application US/09430564  
; Patent No. 6372467  
; GENERAL INFORMATION:  
; APPLICANT: John Blenis  
; APPLICANT: Kay K. Lee-Fruman  
; APPLICANT: Calvin J. Kuo  
; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,  
; FILE REFERENCE: 00246/506002  
; CURRENT APPLICATION NUMBER: US/09/430,564  
; CURRENT FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: 60/106,141  
; PRIOR FILING DATE: 1998-10-29  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-564-8

Query Match 10.7%; Score 49.5; DB 3; Length 85;  
Best Local Similarity 36.6%; Pred. No. 1.1e+02;  
Matches 15; Conservative 6; Mismatches 13; Indels 7; Gaps 2;  
QY 2 EHYRKAGSV---ELPAPSPMPQLPDDT---LEMRVDSK 35  
DB 34 EGFRPSPSLPPTLPLPPLPPPPSTTAPLPIRPSGK 74

RESULT 6  
US-09-430-564-14  
; Sequence 14, Application US/09430564  
; Patent No. 6372467  
; GENERAL INFORMATION:  
; APPLICANT: John Blenis  
; APPLICANT: Kay K. Lee-Fruman  
; APPLICANT: Calvin J. Kuo  
; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,  
; FILE REFERENCE: 00246/506002  
; CURRENT APPLICATION NUMBER: US/09/430,564  
; CURRENT FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: 60/106,141  
; PRIOR FILING DATE: 1998-10-29  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 85  
; TYPE: PRT

```
; ORGANISM: Homo sapiens
US-09-430-564-14

Query Match      10.7%; Score 49.5; DB 3; Length 85;
Best Local Similarity 36.6%; Pred. No. 1.1e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 2; Gaps 2;

QY 2 EHYRKAGSV----ELPAPSPMPQLPPDT---LEMRVRDGSK 35
   | | | | | | | | | | | | | | | | | | | | | |
Db 34 EGRFSPSLPEPTLPLPLPPPPSTTAPLPIRPPSGTK 74
   | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-489-039A-13911
; Sequence 13911, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13911
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13911

Query Match      10.6%; Score 49; DB 4; Length 86;
Best Local Similarity 37.9%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 4; Gaps 1;

QY 13 PAPSMPQLPPD---TLEMRVRDGSKIR 37
   | | | | | | | | | | | | | | | | | | | | | |
Db 51 PRQPQDPDPDEPIKLSHQARSARIR 79
   | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-09-248-796A-16579
; Sequence 16579, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16579
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (75)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknd
US-09-248-796A-16579

Query Match      10.5%; Score 48.5; DB 4; Length 87;
Best Local Similarity 37.8%; Pred. No. 1.5e+02;
Matches 14; Conservative 8; Mismatches 12; Indels 3; Gaps 1;

QY 12 LPAPSPMPQLPPDTLEMRVRDGSKIRNLGLALGRLE 48
   | | | | | | | | | | | | | | | | | | | | | |

; ORGANISM: Homo sapiens
3 LPAPTMGR---EALINRITNETKIQTIALNLDGKKLE 36

RESULT 9
US-09-018-635-37
; Sequence 37, Application US/09018635
; Patent No. 6297356
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Broccoli, Dominique
; APPLICANT: Smogorzewska, Agata
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,635
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: David A. Jackson
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-142 CIPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: TRF2
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-018-635-37

Query Match      10.4%; Score 48; DB 3; Length 45;
Best Local Similarity 44.4%; Pred. No. 73;
Matches 16; Conservative 3; Mismatches 9; Indels 8; Gaps 2;

QY 49 GGSARHVFSGSRAAGKAVSCAEIVKRR---VPGGL 81
   | | | | | | | | | | | | | | | | | | | | | |
Db 5 GGSS-----DGGRAAGRAASRSSGPARRGHPEGL 35
   | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
US-09-912-962-37
; Sequence 37, Application US/09912962
; Patent No. 6586577
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: de Lange, Titia
; APPLICANT: Broccoli, Dominique
; APPLICANT: Smogorzewska, Agata
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
```



; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6477  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-6477

Query Match 10.2%; Score 47; DB 4; Length 59;  
Best Local Similarity 47.4%; Pred. No. 1.4e+02;  
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 5 RKAGSVLPAPSPMPQLPP 23  
| : | | | | | | | | | |  
Db 38 RRPQRRPPPPPPPPPLPP 56

## RESULT 15

US-09-270-767-37413  
; Sequence 37413, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 37413  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-37413

Query Match 10.2%; Score 47; DB 4; Length 60;  
Best Local Similarity 31.7%; Pred. No. 1.4e+02;  
Matches 13; Conservative 6; Mismatches 14; Indels 8; Gaps 2;

Qy 3 HYRKA-----GSVELPAPSPMPQLPPD---TLEMRVRDGSK 35  
| | | : | | : | | : | | : | | :  
Db 2 HYRTAAPSGPAAYELHIPAOPFPCCPGARRSLKYHLPDSSR 42

Search completed: January 26, 2005, 15:36:08  
Job time : 39 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:01 ; Search time 8.2 Seconds  
(without alignments)  
481.084 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_85\_125

Perfect score: 220  
Sequence: 1 TKRLQTEDSWVPASPDTC.....DPLTVRRHVPVWVLLSRDP 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66.5	30.2	1275	2 AD0332	probable membrane
2	57	25.9	1576	2 T29237	hypothetical prote
3	56.5	25.7	160	2 AF3517	lactoylglutathione
4	55	25.0	785	2 G75578	aculeacin A acylas
5	54.5	24.8	3938	2 T42761	Bassoon protein -
6	54.5	24.8	3942	2 T42730	Bassoon protein -
7	54	24.5	4464	2 D87755	protein T21B12.4 [
8	53	24.1	236	2 C69993	ABC transporter (A
9	53	24.1	608	2 AB3562	gtp-binding protei
10	53	24.1	6260	2 T30228	polyketide synthas
11	52.5	23.9	236	2 C82566	hypothetical prote
12	52.5	23.9	301	2 E85902	GTP-binding protei
13	52.5	23.9	301	2 H91057	GTP-binding protei
14	52.5	23.9	411	2 AH0807	probable membrane
15	52.5	23.9	809	2 E98314	probable quinase d
16	52.5	23.9	809	2 A12968	hypothetical prote
17	52	23.6	218	2 E90790	hypothetical prote
18	52	23.6	218	2 H85650	hypothetical prote
19	52	23.6	538	2 H86335	T20H2.2 protein -
20	52	23.6	1247	1 RDECN4	nitrate reductase
21	52	23.6	1247	2 H85702	nitrate reductase
22	52	23.6	1247	2 A99845	nitrate reductase
23	51.5	23.4	349	2 T20202	hypothetical prote
24	51.5	23.4	364	2 S43117	transposase - Lept
25	51.5	23.4	374	2 T33328	hypothetical prote
26	51.5	23.4	771	2 B38252	granulocyte colony
27	51.5	23.4	783	2 JH0329	granulocyte colony
28	51.5	23.4	863	2 C38252	granulocyte colony
29	51	23.2	275	2 F70607	hypothetical prote

ALIGNMENTS

RESULT 1

AD0332  
Probable membrane protein YPO2724 [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AD0332  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AD0332  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1275 <KUR>  
A;Cross-references: UNIPROT:Q8ZD68; GB:AL590842; PIDN:CAC92963.1; PID:g15980702; GSPDB:  
C;Genetics:  
A;Gene: YPO2724

Query Match 30.2%; Score 66.5; DB 2; Length 1275;  
Best Local Similarity 47.1%; Pred. No. 2.2;  
Matches 16; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Qy 3 LRFLQTEDSWVPA-SPDTGLDPLTVRRHVPVWV 35  
Db 806 LMLLQBEDRWVMAQSLSSLDPLTLRQDVLVWYL 839

RESULT 2

T29237  
hypothetical protein F56E3.3 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T29237  
R;Gattung, S.; Wu, X.  
submitted to the EMBL Data Library, November 1995  
A;Description: The sequence of C. elegans cosmid F56E3.  
A;Reference number: Z20592  
A;Accession: T29237  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1576 <GAT>  
A;Cross-references: UNIPROT:Q8IG10; EMBL:U41536; PIDN:AAB52613.1; GSPDB:GN00028; CESP:  
A;Experimental source: strain Bristol N2; clone F56E3  
C;Genetics:  
A;Gene: CESP:F56E3.3  
A;Map position: X  
A;Introns: 23/1; 47/2; 125/3; 162/2; 191/3; 274/2; 296/3; 498/3; 532/1; 651/3; 723/2; 8

Query Match 25.9%; Score 57; DB 2; Length 1576;  
Best Local Similarity 33.3%; Pred. No. 48;

Matches 11; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 8 TEDSVAPSPDGLDPLTVRRHVPVAVLLSRD 40  
Db 1197 TPNSICPGACDWIAPGVERHIPVFLDLNSD 1229

RESULT 3  
AF3517  
lactoylglutathione lyase (EC 4.4.1.5) [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C:Accession: AF3517  
R:DeiVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: ADJ252; PMID:11756688  
A:Accession: AF3517  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-160 <KUR>  
A:Cross-references: UNIPROT:Q8YDWO; UNIPROT:Q8FXN6; GB:AE008918; PIDN:AAJ53305.1; PID:gl  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI10064  
A:Map position: II  
C:Keywords: carbon-sulfur lyase

Query Match 25.7%; Score 56.5; DB 2; Length 160;  
Best Local Similarity 44.4%; Pred. No. 4.5;  
Matches 12; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

QY 15 ASPDTGLDPLTVRRHVPVAVLLSRDP 41  
Db 113 AQLTGRDPVLR---PQTWAFLLHDP 136

RESULT 4  
G75578  
aculeacin A acylase - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: G75578  
R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: G75578  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-785 <WHI>  
A:Cross-references: UNIPROT:Q8YQ04; GB:AE001863; GB:AE001825; NID:G6460670; PIDN:AAF1238  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRA0255  
A:Map position: 2

Query Match 25.0%; Score 55; DB 2; Length 785;  
Best Local Similarity 47.8%; Pred. No. 40;  
Matches 11; Conservative 4; Mismatches 6; Indels 2; Gaps 1;

QY 11 SNV--PASPDGLDPLTVRRHVP 31  
Db 500 AWLANPAAQTGLDPLVGVNAP 522

RESULT 5  
T42761  
Bassoon protein - rat  
N:Alternate names: brain-specific synapse-associated protein

C:Species: Rattus norvegicus (Norway rat)  
C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T42761  
R:Dieck, S.; Sammarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex,  
J. Cell Biol. 142, 499-509, 1998  
A:Title: Bassoon, a novel zinc-finger CAG/glutamine-repeat protein selectively localizes  
A:Reference number: Z22249; MUID:98345363; PMID:9679147  
A:Accession: T42761  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-3938 <DIE>  
A:Cross-references: UNIPROT:O88778; EMBL:Y16563; NID:G3413503; PIDN:CAA76287.1; PID:G341  
A:Experimental source: strain Sprague Dawley; brain  
C:Function:  
A:Description: may be involved in cytomatrix organization at the site of neurotransmitt  
A:Note: component of the presynaptic cytoskeleton  
C:Keywords: coiled coil; zinc finger

Query Match 24.8%; Score 54.5; DB 2; Length 3938;  
Best Local Similarity 35.0%; Pred. No. 2.8e+02;  
Matches 14; Conservative 4; Mismatches 5; Indels 17; Gaps 2;

QY 6 LQTEQWAGRSIGIKRHSMPLRLDACEPSGDPSTVRR 28  
Db 2534 LQTEQWAGRSIGIKRHSMPLRLDACEPSGDPSTVRR 2573

RESULT 6  
T42730  
Bassoon protein - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T42730  
R:Dieck, S.; Sammarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex,  
J. Cell Biol. 142, 499-509, 1998  
A:Title: Bassoon, a novel zinc-finger CAG/glutamine-repeat protein selectively localizes  
A:Reference number: Z22249; MUID:98345363; PMID:9679147  
A:Accession: T42730  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-3942 <DIE>  
A:Cross-references: UNIPROT:O88737; EMBL:Y17034; NID:G3413809; PIDN:CAA76598.1; PID:G341  
A:Experimental source: strain 129 SVJ  
C:Genetics:  
A:Map position: 9F1  
A:Introns: 72/2; 208/3; 505/3; 675/3; 2889/3; 3582/1; 3851/3; 3886/1; 3930/1  
A:Note: bassoon  
C:Function:  
A:Description: may be involved in cytomatrix organization at the site of neurotransmitt  
A:Note: component of the presynaptic cytoskeleton  
C:Keywords: coiled coil; zinc finger

Query Match 24.8%; Score 54.5; DB 2; Length 3942;  
Best Local Similarity 35.0%; Pred. No. 2.8e+02;  
Matches 14; Conservative 4; Mismatches 5; Indels 17; Gaps 2;

QY 6 LQTEQWAGRSIGIKRHSMPLRLDACEPSGDPSTVRR 28  
Db 2548 LQTEQWAGRSIGIKRHSMPLRLDACEPSGDPSTVRR 2587

RESULT 7  
D87755  
Protein T21E12.4 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: D87755  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and



A;Accession: D87755  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-4464 <STO>  
A;Cross-references: GB:chr\_1; PIDN:AAB37647.1; PID:gl1703597; GSPDB:GN00019; CESP:T21E12.1  
C;Genetics:  
A;Gene: T21E12.4  
A;Map position: 1  
C;Superfamily: dynein heavy chain, cytosolic

Query Match 24.5%; Score 54; DB 2; Length 4464;  
Best Local Similarity 37.8%; Pred. No. 20;  
Matches 17; Conservative 3; Mismatches 11; Indels 14; Gaps 3;

QY 5 FLQTEDSWVPAS-----PDTGLDPLTVR--RHVPAYMV 35  
| : | | | | | | | | | | | | | | | | | | | |  
Db 2530 FYRTSDHSWSLERIQFVGACNPPTDPGRHPMTSRFLRHVPYV 2574  
| : | | | | | | | | | | | | | | | | | | | |

RESULT 8  
ABC transporter (ATP-binding protein) homolog ythP - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: C69993  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta  
C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerman, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: C69993  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-236 <KUN>  
A;Cross-references: UNIPROT:O34977; GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB14979.  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: ythP  
C;Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology  
F;21-201/Domain: ATP-binding cassette homology <ABC>  
F;38-45/Region: nucleotide-binding motif A (P-loop)

Query Match 24.1%; Score 53; DB 2; Length 236;  
Best Local Similarity 42.3%; Pred. No. 20;  
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 5 FLQTEDSWVPASPDPTGLDPLTVRRHV 30  
| : | | | | | | | | | | | | | | | | | | | |  
Db 140 FLKPDVMYVDEPPFGLDPISTRFV 165  
| : | | | | | | | | | | | | | | | | | | | |

RESULT 9  
gtp-binding protein typA/bipA [imported] - Brucella melitensis (strain 16M)  
C;Species: Brucella melitensis  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C;Accession: AB3562  
R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688  
A;Accession: AB3562  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-608 <KUR>  
A;Cross-references: UNIPROT:Q8YCV9; GB:AE008918; PIDN:AAL53661.1; PID:gl7984580; GSPDB:  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEI10419  
A;Map position: II  
C;Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homo

Query Match 24.1%; Score 53; DB 2; Length 608;  
Best Local Similarity 33.3%; Pred. No. 56;  
Matches 16; Conservative 5; Mismatches 11; Indels 16; Gaps 3;

QY 10 DSWVPASP---DPTGLDPL--TVRRHVPA-----VWVLSRDP 41  
| : | | | | | | | | | | | | | | | | | | | |  
Db 167 NGWMLSPGPKDQGLAPLFDLVLKHVPAPKVAEGPFRMIGTILEADP 214  
| : | | | | | | | | | | | | | | | | | | | |

RESULT 10  
T30228  
polyketide synthase - Streptomyces hygroscopicus  
C;Species: Streptomyces hygroscopicus  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T30228  
R;Aparicio, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Stau  
Gene 169, 9-16, 1996  
A;Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hy  
A;Reference number: Z20782; MUID:96186896; PMID:8635756  
A;Accession: T30228  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-6260 <AP>  
A;Cross-references: UNIPROT:Q54299; EMBL:X86780; NID:g987088; PID:g987102; PIDN:CAA6046  
C;Genetics:  
A;Gene: rapC  
C;Keywords: carrier protein  
F;55-451/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>  
F;1536-1607/Domain: acyl carrier protein homology <ACP1>  
F;1651-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>  
F;3123-3194/Domain: acyl carrier protein homology <ACP2>  
F;3238-3632/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>  
F;3725-3999/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>  
F;5114-5185/Domain: acyl carrier protein homology <ACP3>  
F;5242-5638/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>  
F;5759-6044/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>  
F;6135-6206/Domain: acyl carrier protein homology <ACP4>

Query Match 24.1%; Score 53; DB 2; Length 6260;  
Best Local Similarity 39.3%; Pred. No. 7.2e+02;  
Matches 11; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 14 PASPDPTGLDPLTVRRHVPAVWVLSRDP 41  
| : | | | | | | | | | | | | | | | | | | | |  
Db 5653 PAPPVTPVTPVTNPBPGLPWLVAQSP 5680  
| : | | | | | | | | | | | | | | | | | | | |

RESULT 11  
C82566  
hypothetical protein XP2359 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: C82566  
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: AB3515; MUID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: C82566  
A;Status: preliminary  
A;Molecule type: DNA



Db 396 PGNPDITKLPPEGQTYTRSTPNVWMSYDP 426

Search completed: January 26, 2005, 15:07:33  
Job time : 9.2 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:40 ; Search time 43.9186 Seconds  
(without alignments)  
537.138 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_85\_125

Perfect score: 220

Sequence: 1 TKLRFLOTEDSWVPASPDGTG.....DPLTVRRHVPAWVLLSRDP 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220	100.0	163	2 Q8N5L8	Q8N5L8 homo sapien
2	216	98.2	163	2 Q99JH1	Q99JH1 mus musculus
3	216	98.2	163	2 Q6CYX4	Q6CYX4 mus musculus
4	125	56.8	224	2 Q6DGS1	Q6DGS1 brachydanio
5	69	31.4	199	2 Q91WE3	Q91WE3 mus musculus
6	68	30.9	199	2 Q9NX88	Q9NX88 homo sapien
7	68	30.9	199	2 Q9BUL9	Q9BUL9 homo sapien
8	67	30.5	286	2 Q96198	Q96198 homo sapien
9	67	30.5	553	1 Q324 HUMAN	Q324 HUMAN
10	66.5	30.2	1192	2 Q8CJD8	Q8CJD8 yersinia pe
11	66.5	30.2	1192	2 AAS62645	AAS62645 yersinia pe
12	66.5	30.2	1275	2 Q8ZD68	Q8ZD68 yersinia pe
13	64	29.1	206	2 Q6NNC7	Q6NNC7 drosophila
14	64	29.1	206	2 Q9V9B6	Q9V9B6 drosophila
15	64	29.1	206	2 AAR96155	AAR96155 drosophila
16	63.5	28.9	436	2 Q82GR6	Q82GR6 streptomyc
17	60.5	27.5	490	2 Q6MZA1	Q6MZA1 corynebacte
18	60.5	27.5	490	2 CAF20771	CAF20771 corynebac
19	59.5	27.0	479	2 Q8LMV2	Q8LMV2 oryza sativ
20	59	26.8	484	2 Q7ZVT2	Q7ZVT2 xenopus lae
21	58.5	26.6	354	2 Q72B09	Q72B09 desulfovibr
22	58.5	26.6	354	2 AAS63307	AAS63307 desulfovibr
23	58	26.4	460	2 Q8HMI6	Q8HMI6 carapace ber
24	58	26.4	608	2 Q6GGR7	Q6GGR7 bartonella
25	58	26.4	608	2 Q6G4X9	Q6G4X9 bartonella
26	58	26.4	4557	2 Q8CJN6	Q8CJN6 streptomyc
27	57.5	26.1	321	2 Q89EB5	Q89EB5 bradyrhizob
28	57.5	26.1	690	2 Q89TL8	Q89TL8 bradyrhizob
29	57	25.9	215	2 Q7PTH5	Q7PTH5 anopheles g
30	57	25.9	262	2 Q7XT90	Q7XT90 oryza sativ
31	57	25.9	606	2 Q72PW0	Q72PW0 leptospira

32	57	25.9	606	2 Q8F6D0	Q8F6D0 leptospira
33	57	25.9	606	2 AAS70926	AAS70926 leptospir
34	57	25.9	793	2 Q9NDT3	Q9NDT3 caenorhabdi
35	57	25.9	1576	2 Q81G10	Q81G10 caenorhabdi
36	57	25.9	1595	2 Q20888	Q20888 caenorhabdi
37	57	25.9	1610	2 Q8MQ20	Q8MQ20 caenorhabdi
38	56.5	25.7	160	2 Q8FXN6	Q8FXN6 bruceella su
39	56.5	25.7	160	2 Q8YDW0	Q8YDW0 bruceella su
40	56.5	25.7	219	2 Q6V273	Q6V273 symbiont ba
41	56.5	25.7	219	2 AAQ91803	AAQ91803 symbiont
42	56.5	25.7	866	2 Q8S2P1	Q8S2P1 oryza sativ
43	56.5	25.7	1460	1 PCX1 MOUSE	PCX1 mus musculus
44	56.5	25.7	1751	2 Q6ZQ41	Q6ZQ41 mus musculu
45	56.5	25.7	1751	2 BAC98031	BAC98031 mus muscu

ALIGNMENTS

RESULT 1

ID	Q8N5L8	PRELIMINARY;	PRT;	163 AA.
AC	Q8N5L8;			
DT	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Chromosome 9 open reading frame 23 protein.			
GN	Name=C9orf23;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OC	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Raha S.S., Loquellano N.A., Toshiyuki S., Carninci P., Prange C.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Guimond J., Schmutz J., Myers R.M., Butterfield V.S.,			
RA	Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RA	Strausberg R.;			
RA	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; BC032136; AAH32136.1; -			
DR	Genew; HGNC:19909; C9orf23.			
DR	InterPro; IPR011574; Alba_DUF78.			
DR	ProDom; PD010497; Alba_DUF78; 1.			
SQ	SEQUENCE 163 AA; 17631 MW; 442C8727191A0BCE CRC64;			

Query Match 100.0%; Score 220; DB 2; Length 163;  
Best Local Similarity 100.0%; Pred. No. 3.7e-21; Indels 0; Gaps 0;  
Matches 41; Conservative 0; Mismatches 0;  
1 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAWVLLSRDP 41  
|||||

Db 85 TKLRFLOTDSWVPSPDTGLDPLTVRRHVPAVWLLSRDP 125

RESULT 2

Q99JH1 PRELIMINARY; PRT; 163 AA.

AC Q99JH1 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein (RIKEN cDNA 2810432D09).

DE Mus musculus (Mouse).

GN Name=2810432D09Rik;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N-3;

RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.

RC MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Whiting M., Madan A., Young A.C., Shchepochko Y., Bouffard G.G., Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;

RA "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N-3;

RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.

RA Strausberg R.J.

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC293897; CAC34588.1; -

DR EMBL; BC047068; AAH47068.1; -

DR MGD; MGI:1917211; 2810432D09Rik.

DR InterPro; IPR011574; Alba\_DUF78.

DR ProDom; PD010497; Alba\_DUF78; 1.

KW Hypothetical protein.

SQ SEQUENCE 163 AA; 17675 MW; 87675201AF87F5B6 CRC64;

Query Match 98.2%; Score 216; DB 2; Length 163;

Best Local Similarity 97.6%; Pred. No. 1.3e-20;

Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TKLRFLOTDSWVPSPDTGLDPLTVRRHVPAVWLLSRDP 41

Db 85 TKLRFLOTDSWVPSPDTGLDPLTVRRHVPAVWLLSRDP 125

RESULT 3

Q9CYX4 PRELIMINARY; PRT; 163 AA.

AC Q9CYX4 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810432D09 product:hypothetical protein, full insert sequence.

GN Name=2810432D09Rik;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RC MEDLINE=99279253; PubMed=10349636;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning."

RL Meth. Enzymol. 303:19-44(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RC MEDLINE=21085660; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RC MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."

RL Genome Res. 10:1617-1630(2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RC MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multiplex capillary sequencer."

RL Genome Res. 10:1757-1771(2000).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T., Sojabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Ito T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; AK013222; BAB28723.1; -.
DR MGd; MGI:1917211; 2810432D09Rik.
DR InterPro; IPRO11574; Alba DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
KW Hypothetical protein.
SQ SEQUENCE 163 AA; 17648 MW; DC89421P4F617429 CRC64;

Query Match          98.2%; Score 216; DB 2; Length 163;
Best Local Similarity 97.6%; Pred. No. 1.3e-20;
Matches 40; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 TKLRFLOTEDSWPASPDTGLDLPTVRRHVPAYVWLLSRDP 41
   |||||
Db 85 TKLRFLOTEDSWPASPDTGLDLPTVRRHVPAYVWLLSRDP 125

RESULT 4
Q6DGS1 .PRELIMINARY; PRT; 224 AA.
AC Q6DGS1;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RN [2]
RP "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).

```

```

(7)
RN  SEQUENCE FROM N.A.
RP  STRAIN=C57BL/6J;
RX  MEDLINE=20530913; PubMed=11076861;
RA  Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA  Konno H., Akiyama J., Nishi K., Kitagawa T., Tashiro H., Itoh M.,
RA  Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA  Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA  Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA  Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA  Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT  "RIKEN integrated sequence analysis (RISA) system-384-format
RT  sequencing pipeline with 384 multicapillary sequencer.";
RL  Genome Res. 10:1757-1771 (2000).
RN  [8]
RP  SEQUENCE FROM N.A.
RA  STRAIN=C57BL/6J;
RA  Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA  Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA  Hayaehida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA  Hori F., Imotani K., Ishii Y., Itoh M., Kigawa I., Kasukawa T.,
RA  Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA  Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA  Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA  Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA  Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA  Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA  Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL  Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC016085; AH16085.1; -.
DR  EMBL; AK082822; BAC38636.1; -.
DR  MGD; MGI:2143151; AI851185.
DR  InterPro; IPR011574; Alba_DUF78.
DR  ProDom; PD010497; Alba_DUF78; 1.
KW  Hypothetical protein.
SQ  SEQUENCE 199 AA; 21037 MW; B9DE105BE3293D64 CRC64;

```

```

Query Match          31.4%; Score 69; DB 2; Length 199;
Best Local Similarity 33.3%; Pred. No. 0.64;
Matches 16; Conservative 11; Mismatches 13; Indels 8; Gaps 2;

```

```

OY  1 TKLRFLOQTESW--VPASPDGTGLDP-----LTVRRHVPVAVVLLSRD 40
DB  95 TRLRYSRVREVMQSLPPGPTQGTPGPAASLSVLKNVPSLAILLSKD 142

```

```

RESULT 6
Q9NX88 ID Q9NX88 PRELIMINARY; PRT; 199 AA.
AC Q9NX88;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ20374.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
RA Hiro M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
RA Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000381; BAA91128.1; -.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 199 AA; 20660 MW; 855494F0366B2291 CRC64;

```

```

Query Match          30.9%; Score 68; DB 2; Length 199;
Best Local Similarity 33.3%; Pred. No. 0.87;
Matches 16; Conservative 10; Mismatches 14; Indels 8; Gaps 2;

```

```

OY  1 TKLRFLOQTESW--VPASPDGTGLDP-----LTVRRHVPVAVVLLSRD 40
DB  95 TRLRYSRVREVMQSLPPGPTQGTPGPAASLSVLKNVPSLAILLSKD 142

```

RESULT 7

Q9BUL9 ID Q9BUL9 PRELIMINARY; PRT; 199 AA.

AC Q9BUL9;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE RPP25 protein (RNase P protein subunit p25).

DE Name=RPP25;

GN Name=RPP25;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA TISSUE=Brain, and Skin;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,

RA Krzywinski M.I., Skalska J., Smalut D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Skin;

RA Strausberg R.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Strausberg R.;

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=21998191; PubMed=12003489;

RA Guerrier-Takada C., Eder P.S., Gopalan V., Altman S.;

RT "Purification and characterization of Rpp25, an RNA-binding protein

RT subunit of human ribonuclease P.";

RT RNA 8:290-295 (2002).

DR EMBL; BC002497; AAH02497.1; -.

DR EMBL; BC007270; AAH07270.1; -.

DR EMBL; AY034074; AAK54443.1; -.

DR Inact; Q9BUL9;

DR InterPro; IPR011574; Alba\_DUF78.

DR ProDom; PD010497; Alba\_DUF78; 1.

SQ SEQUENCE 199 AA; 20632 MW; 9A4494F0297B2A81 CRC64;

Query Match 30.9%; Score 68; DB 2; Length 199;

Best Local Similarity 33.3%; Pred. No. 0.87;

Matches 16; Conservative 10; Mismatches 14; Indels 8; Gaps 2;

```

OY  1 TKLRFLOQTESW--VPASPDGTGLDP-----LTVRRHVPVAVVLLSRD 40
DB  95 TRLRYSRVREVMQSLPPGPTQGTPGPAASLSVLKNVPSLAILLSKD 142

```



Db 95 TRLRYRSVREVWOSLPDPTGQTPGEPAAASLVKXVPGLAILLKSD 142

RESULT 8

Q96I98 PRELIMINARY; PRT; 286 AA.

AC Q96I98; 2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE ZNF324 protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Uterus;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieff F.,

RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Kzywinski M.I., Skalka U., Smaluk D.E., Scherch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]

RP SEQUENCE FROM N.A.

RC TISSUE=Uterus;

RA Strausberg R.;

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC007717; AA007717.1; -.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001909; KRAB.

DR InterPro; IPR001005; MYB DNA binding.

DR InterPro; IPR007087; Znf\_C2H2.

DR Pfam; PF01352; KRAB; 1.

DR SMART; SM00349; KRAB; 1.

DR PROSITE; PS00355; Znf\_C2H2; 1.

DR PROSITE; PS00805; KRAB; 1.

DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 9.

DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 9.

KW DNA-binding; Metal-binding; Nuclear protein; Repeat;

KW Transcription regulation; Zinc-finger.

FT DOMAIN 1 72 KRAB.

FT ZN\_FING 257 279 C2H2-type 1.

FT ZN\_FING 285 307 C2H2-type 2.

FT ZN\_FING 313 335 C2H2-type 3.

FT ZN\_FING 341 363 C2H2-type 4.

FT ZN\_FING 369 391 C2H2-type 5.

FT ZN\_FING 397 419 C2H2-type 6.

FT ZN\_FING 425 447 C2H2-type 7.

FT ZN\_FING 453 475 C2H2-type 8.

FT ZN\_FING 481 503 C2H2-type 9.

SQ SEQUENCE 553 AA; 61104 MW; BCB5C469A13108F CRC64;

Query Match 30.5%; Score 67; DB 2; Length 286;

Best Local Similarity 46.4%; Pred. No. 1.8;

Matches 13; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 9 EDSWVPASPTGLDPLTVRRHVPVAVWL 36

Db 57 EEPWVPSGDTTLRTTYRRRNPQSWSL 84

RESULT 9

Z324\_HUMAN

ID 2324 HUMAN

AC 075467;

QY 9 EDSWVPASPTGLDPLTVRRHVPVAVWL 36

Db 57 EEPWVPSGDTTLRTTYRRRNPQSWSL 84

RESULT 10

Q8CLD8

ID Q8CLD8 PRELIMINARY; PRT; 1192 AA.

AC Q8CLD8; Q74SW5;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical (Putative membrane protein).

GN Name=icmF2; OrderedLocusNames=yp2442, y1553;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Zinc finger protein 324 (Zinc finger protein ZF5128).

GN Name=ZNF324; (Human).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Kim B.W., Cho H.M., Jun D.Y., Kim Y.H.;

RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: May function as a transcription factor.

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein family.

CC -!- SIMILARITY: Contains 9 C2H2-type zinc fingers.

CC -!- SIMILARITY: Contains 1 KRAB domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/ebis-sib.ch/>).

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

CC -----

DR EMBL; AF060503; AAC33716.1; -.

DR HSSP; P08047; ISP2.

DR Genew; HGNC:14096; ZNF324.

DR InterPro; IPR001909; KRAB.

DR InterPro; IPR007087; Znf\_C2H2.

DR Pfam; PF01352; KRAB; 1.

DR Pfam; PF00096; zf-C2H2; 9.

DR ProDom; PD000003; Znf\_C2H2; 5.

DR SMART; SM00349; KRAB; 1.

DR SMART; SM00355; Znf\_C2H2; 9.

DR PROSITE; PS00805; KRAB; 1.

DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 9.

DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 9.

KW DNA-binding; Metal-binding; Nuclear protein; Repeat;

KW Transcription regulation; Zinc-finger.

FT DOMAIN 1 72 KRAB.

FT ZN\_FING 257 279 C2H2-type 1.

FT ZN\_FING 285 307 C2H2-type 2.

FT ZN\_FING 313 335 C2H2-type 3.

FT ZN\_FING 341 363 C2H2-type 4.

FT ZN\_FING 369 391 C2H2-type 5.

FT ZN\_FING 397 419 C2H2-type 6.

FT ZN\_FING 425 447 C2H2-type 7.

FT ZN\_FING 453 475 C2H2-type 8.

FT ZN\_FING 481 503 C2H2-type 9.

SQ SEQUENCE 553 AA; 61104 MW; BCB5C469A13108F CRC64;

Query Match 30.5%; Score 67; DB 1; Length 553;

Best Local Similarity 46.4%; Pred. No. 3.8;

Matches 13; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 9 EDSWVPASPTGLDPLTVRRHVPVAVWL 36

Db 57 EEPWVPSGDTTLRTTYRRRNPQSWSL 84

OS	Yersinia pestis.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Yersinia.
OX	NCBI_TaxID=632;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=KIM5 / Biovar Mediaevalis;
RX	MEDLINE=22137863; PubMed=12142430;
RA	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA	Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA	Parry R.D.;
RL	"Genome sequence of Yersinia pestis KIM.";
RL	J. Bacteriol. 184:4601-4611(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=91001 / Biovar Mediaevalis;
RA	Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA	Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA	Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA	Yang R.;
RL	Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AE0113758; AAC85122.1; --
DR	ENML; AE017136; AAS62645.1; --
DR	InterPro; IPR010623; DUF1215.
DR	InterPro; IPR009612; ImcF-related.
DR	Pfam; PF06744; DUF1215; 1.
DR	Pfam; PF06761; ImcF-related; 1.
KW	Hypothetical protein.
SQ	SEQUENCE 1192 AA; 135060 MW; C53ECD6EAD1D8804F CRC64;
Query Match 30.2%; Score 66.5; DB 2; Length 1192;	
Best Local Similarity 47.1%; Pred. No. 11;	
Matches 16; Conservative 6; Mismatches 11; Indels 1; Gaps	
Qy	3 LRFLQTESWVPA-SPDTGLDLTVRRHVPVVW 35
Db	723 LMLQEERWVMAQSLSGLDLTLRDVLVMYL 756
RESULT 11	
AAS62645	
ID	AAS62645 PRELIMINARY; PRT; 1192 AA.
AC	AAS62645;
DC	24-MAR-2004 (TrEMBLrel. 27, Created)
DT	24-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT	04-NAR-2004 (TrEMBLrel. 27, Last annotation update)
DE	Putative membrane protein.
GN	ICME2 OR YP2442.
OS	Yersinia pestis.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Yersinia.
OX	NCBI_TaxID=632;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=91001 / Biovar Mediaevalis;
RA	Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA	Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA	Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA	Yang R.;
RL	Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AE017136; AAS62645.1; --
SQ	SEQUENCE 1192 AA; 135060 MW; C53ECD6EAD1D8804F CRC64;
Query Match 30.2%; Score 66.5; DB 2; Length 1192;	
Best Local Similarity 47.1%; Pred. No. 11;	
Matches 16; Conservative 6; Mismatches 11; Indels 1; Gaps	
Qy	3 LRFLQTESWVPA-SPDTGLDLTVRRHVPVVW 35
Db	723 LMLQEERWVMAQSLSGLDLTLRDVLVMYL 756



Search completed: January 26, 2005, 15:06:29  
Job time : 45.9186 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:04 ; Search time 40.861 Seconds  
(without alignments)  
359.949 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_85\_125

Perfect score: 220  
Sequence: 1 TKLRFQTEDEWVPASPDGTG.....DPLTVRRHVPVAVVLLSRDP 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseqp\_23Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220	100.0	163	4	AAG89339 Human sec
2	220	100.0	163	6	ABU00014 Human nov
3	220	100.0	219	8	ADM90988 Human pha
4	194	88.2	163	2	AA36104 Extended
5	194	88.2	163	8	ADP19412
6	68	30.9	211	4	AAM38973 Human pol
7	67	30.5	553	8	ADQ17615 Human sof
8	67	30.5	558	8	ADN99926 Novel hum
9	64	29.1	205	4	ABB64410 Drosophil
10	61.5	28.0	96	5	ABP05282 Human ORF
11	61.5	28.0	586	7	ABO64395 Kiebsiell
12	58	26.4	1037	7	ADJ71002 Human hea
13	57	25.9	408	7	ADC36134 Weed cont
14	56.5	25.7	4083	8	ADO13847 Protein e
15	56	25.5	264	4	ABG21326 Novel hum
16	56	25.5	266	2	AAR43576 Bovine ad
17	56	25.5	298	3	AAG51877 Arabidops
18	56	25.5	298	3	AAG16840 Arabidops
19	56	25.5	314	3	AAG51848 Arabidops
20	56	25.5	314	3	AAG16839 Arabidops
21	56	25.5	314	3	AAG51876 Arabidops
22	56	25.5	350	3	AAG16838 Arabidops
23	56	25.5	350	3	AAG51875 Arabidops
24	56	25.5	350	3	AAG51847 Arabidops
25	56	25.5	350	8	ADN72883 Thale cre

26	56	25.5	350	8	ADN72183	Adn72183 Thale cre
27	56	25.5	356	3	AAG51846	Aag51846 Arabidops
28	55.5	25.2	402	6	ABU24856	Abu24856 Protein e
29	55.5	25.2	1379	4	AAM78555	Aam78555 Human pro
30	55.5	25.2	1397	4	AAM79539	Aam79539 Human pro
31	55.5	25.2	1446	5	AAO18739	Aao18739 Human NOV
32	55.5	25.2	1451	6	ABJ39112	Abj39112 Molecule
33	55	25.0	785	7	ADC03509	Adc03509 Deinococc
34	54.5	24.8	98	4	AAM53880	Aam53880 Propionib
35	54.5	24.8	98	6	ABM50399	Abm50399 Propionib
36	54.5	24.8	245	6	ABU44710	Abu44710 Protein e
37	54.5	24.8	1124	7	ADB70297	Adb70297 C. neofo
38	54	24.5	72	8	ABO55760	Abos5760 Human gen
39	54	24.5	346	5	ABG97461	Abg97461 S. kaniha
40	54	24.5	1246	6	ABU32194	Abu32194 Protein e
41	54	24.5	1282	7	ABO63127	Abos63127 Klebsiell
42	53.5	24.3	144	3	ABB23464	Abb23464 Wheat LLS
43	53	24.1	185	4	AAM40741	Aam40741 Human pol
44	53	24.1	215	3	AAG33138	Aag33138 Zea maye
45	53	24.1	239	3	AAG33137	Aag33137 Zea maye

ALIGNMENTS

RESULT 1  
AAG89339  
ID AAG89339 standard; protein; 163 AA.  
XX AAG89339;  
XX  
DT 11-SRP-2001 (first entry)  
XX  
DE Humán secreted protein, SEQ ID NO: 459.  
XX  
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
KW GENSET.  
XX  
OS Homo sapiens.  
XX  
PN WO200142451-A2.  
XX  
PD 14-JUN-2001.  
XX  
PF 07-DEC-2000; 2000WO-IB001938.  
XX  
PR 08-DEC-1999; 99US-0169629P.  
PR 06-MAR-2000; 2000US-0187470P.  
(GEST ) GENSET.  
Dumas Milne Edwards J, Bougueleret L, Jobert S;  
WPI; 2001-367870/38.  
N-PSDB; AAM64942.  
Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.  
Claim 21; Page 910-911; 921pp; English.  
The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar nucleic acid sequences in samples, and hence to determine which  
CC patients may be in need of restorative therapy. The GENSET polypeptides  
CC may be used as antigens in the production of antibodies and in assays to  
CC identify modulators (agonists and antagonists) of GENSET polypeptide  
CC expression and activity. The present sequence is a GENSET polypeptide of  
CC the invention  
XX  
SQ Sequence 163 AA;

Query Match 100.0%; Score 220; DB 4; Length 163;  
Best Local Similarity 100.0%; Pred. No. 8.2e-23;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLRFLOQEDSWVPASPTGLDPLTVRRHVPVAVVLLSRDP 41  
|||||  
Db 85 TKLRFLOQEDSWVPASPTGLDPLTVRRHVPVAVVLLSRDP 125

RESULT 2  
ABU00014  
ID ABU00014 standard; protein; 163 AA.

XX AC ABU00014;

XX DT 17-JAN-2003 (first entry)

XX DE Human novel polypeptide #107.

XX KW Human; genetic disorder; gene mapping; medical imaging; cancer;  
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;  
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;  
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;  
KW fungal infection; bacterial infection; autoimmune disease; diabetes;  
KW atopic dermatitis.

XX OS Homo sapiens.

XX PN WO200274961-A1.

XX PD 26-SEP-2002.

XX PF 14-MAR-2002; 2002WO-US005109.

XX PR 15-MAR-2001; 2001US-00810173.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX DR WPI; 2003-040556/03.

XX DR N-PSDB; ABX05092.

XX PT New isolated polypeptides and polynucleotides, useful for preventing,  
PT treating or ameliorating medical conditions, such as cancer,  
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative  
PT disorders, and infections.

XX PS Claim 9; SEQ ID NO 633; 235pp; English.

XX CC The invention relates to human polynucleotides and the polypeptides they  
CC encode. The polynucleotides and polypeptides are useful in diagnostics,  
CC forensics, gene mapping, medical imaging, identification of mutations  
CC responsible for genetic disorders or other traits, assessing biodiversity  
CC and producing many other types of data and products dependent on DNA and  
CC amino acid sequences. They are also useful for preventing, treating or  
CC ameliorating medical conditions, such as cancer, neurodegenerative  
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell  
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,  
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or  
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).

CC Sequences ABG99888-ABG9989 and ABU00010-ABU00433 represent human  
CC polypeptides of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied by the European Patent Office

XX SQ Sequence 163 AA;

Query Match 100.0%; Score 220; DB 6; Length 163;  
Best Local Similarity 100.0%; Pred. No. 8.2e-23;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLRFLOQEDSWVPASPTGLDPLTVRRHVPVAVVLLSRDP 41  
|||||  
Db 85 TKLRFLOQEDSWVPASPTGLDPLTVRRHVPVAVVLLSRDP 125

RESULT 3  
ADM90988

ID ADM90988 standard; protein; 219 AA.

XX AC ADM90988;

XX DT 03-JUN-2004 (first entry)

XX DE Human pharmaceutically useful protein SeqID 381.

XX KW human; cancer; haematopoiesis; thrombosis; anaemia;  
KW cardiovascular disorder; ischaemic heart disease;  
KW acute myocardial infarction; respiratory disease; asthma; pneumonia;  
KW cystic fibrosis; chronic renal failure; glomerulopathy;  
KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;  
KW HIV infection; systemic lupus erythematosus; endocrine system;  
KW diabetes mellitus; epilepsy; Alzheimer's disease;  
KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;  
KW fungal; parasitic; viral infection; cytostatic; anticoagulant;  
KW thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;  
KW antiinflammatory; nephrotropic; antiulcer; hepatotropic;  
KW immunosuppressive; antiallergic; dermatological; antirheumatic;  
KW antiarthritic; antidiabetic; genetic; gene therapy; vaccine;  
KW antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;  
KW gene therapy; vaccine.

XX OS Homo sapiens.

XX PN WO2004020595-A2.

XX PD 11-MAR-2004.

XX PF 28-AUG-2003; 2003WO-US027107.

XX PR 29-AUG-2002; 2002US-0406576P.

XX PR 29-AUG-2002; 2002US-0406611P.

XX PR 29-AUG-2002; 2002US-0406612P.

XX PR 29-AUG-2002; 2002US-0406616P.

XX PR 29-AUG-2002; 2002US-0406640P.

XX PR 29-AUG-2002; 2002US-0406655P.

XX PR 17-SEP-2002; 2002US-0406666P.

XX PR 17-SEP-2002; 2002US-0410946P.

XX PR 17-SEP-2002; 2002US-0410951P.

XX PR 17-SEP-2002; 2002US-0410953P.

XX PR 17-SEP-2002; 2002US-0410957P.

XX PR 17-SEP-2002; 2002US-0410960P.

XX PR 17-SEP-2002; 2002US-0411019P.

XX PR 17-SEP-2002; 2002US-0411022P.

XX PR 17-SEP-2002; 2002US-0411024P.

XX PR 17-SEP-2002; 2002US-0411037P.

XX PR 17-SEP-2002; 2002US-0411046P.

XX PR 17-SEP-2002; 2002US-0411052P.

XX PR 17-SEP-2002; 2002US-0411082P.

XX PR 17-SEP-2002; 2002US-0411111P.

XX PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.  
 XX (DNAF-) DNAFORM KK.  
 PI Williams LT, Chu K, Lee E, Heestir K;  
 XX  
 XX WPI; 2004-257410/24.  
 DR N-PSDB; ADM90779, ADM91197.  
 XX  
 XX New human polynucleotides and polypeptides, useful for diagnosing,  
 PT preventing and treating proliferative disorders, immune disorders,  
 PT cardiovascular disorders, or bacterial, fungal, parasitic and viral  
 PT diseases.  
 XX  
 PS Claim 1; SEQ ID NO 381; 254pp; English.  
 XX  
 CC This invention relates to novel isolated human polynucleotides and the  
 CC encoded proteins thereof. Specifically, it refers to proteases, kinases,  
 CC phosphatases, secreted and transmembrane proteins, as well as the derived  
 CC peptide fragments, which can be used to develop antibodies and screen for  
 CC small molecule agonists and antagonists that can modulate their  
 CC activities. The present invention describes polypeptides,  
 CC polynucleotides, vectors and host cells useful for diagnosing, preventing  
 CC and treating proliferative disorders, e.g. cancer, disorders of  
 CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,  
 CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory  
 CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the  
 CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,  
 CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune  
 CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders  
 CC of the endocrine system, e.g. diabetes mellitus, central nervous system  
 CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral  
 CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,  
 CC parasitic and viral diseases. Accordingly, they exhibit many various  
 CC activities including cytostatic, anticoagulant, thrombolytic,  
 CC antianaemic, cardiant, vasotropic, antiasthmatic, antiinflammatory,  
 CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,  
 CC dermatological, antineumatic, antiarthritic, antidiabetic,  
 CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,  
 CC fungicide, antiparasitic and virucidal, such that these polynucleotides  
 CC can be used for gene therapy purposes and the development of appropriate  
 CC vaccines. This polypeptide is a human protein of the invention.  
 XX  
 SQ Sequence 219 AA;  
 Query Match 100.0%; Score 220; DB 8; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-22;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVLLSRDP 41  
 |||||  
 Db 141 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVLLSRDP 181  
 |||||  
 RESULT 4  
 AAY36104  
 ID AAY36104 standard; protein; 163 AA.  
 XX  
 AC AAY36104;  
 XX  
 DT 13-SEP-1999 (first entry)  
 XX  
 DE Extended human secreted protein sequence, SEQ ID NO. 489.  
 XX  
 KW Secreted protein; human; cytokine; cellular proliferation; cell movement;  
 KW cellular differentiation; immune system regulator; anti-inflammatory;  
 KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;  
 KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;  
 KW genetic disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9931236-A2.  
 XX

PD 24-JUN-1999.  
 XX  
 PF 17-DEC-1998; 98WO-IB002122.  
 XX  
 PR 17-DEC-1997; 97US-0069957P.  
 PR 09-FEB-1998; 98US-0074121P.  
 PR 13-APR-1998; 98US-0081563P.  
 PR 10-AUG-1998; 98US-0096116P.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 XI Bougueleret L, Duclert A, Dumas Milne Edwards J;  
 PI  
 XX WPI; 1999-385906/32.  
 DR N-PSDB; AAX97788.  
 XX  
 PT New isolated human secreted proteins.  
 XX  
 PS Claim 9; Page 423-424; 516pp; English.  
 XX  
 CC This sequence is encoded by an extended human secreted protein coding  
 CC sequence of the invention. The secreted proteins can be used in treating  
 CC or controlling a variety of human conditions. The secreted proteins may  
 CC act as cytokines or may affect cellular proliferation or differentiation  
 CC or may act as immune system regulators, haematopoiesis regulators, tissue  
 CC growth regulators, regulators of reproductive hormones or cell movement  
 CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or  
 CC tumour inhibition activity. The DNAs can be used in forensic procedures  
 CC to identify individuals or in diagnostic procedures to identify  
 CC individuals having genetic diseases resulting from abnormal expression of  
 CC the genes corresponding to the extended cDNAs. They are also useful for  
 CC constructing a high resolution map of the human chromosomes. They can  
 CC also be used for gene therapy to control or treat genetic diseases  
 XX  
 SQ Sequence 163 AA;  
 Query Match 88.2%; Score 194; DB 2; Length 163;  
 Best Local Similarity 87.8%; Pred. No. 3.9e-19;  
 Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVLLSRDP 41  
 |||||  
 Db 85 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVLLSRDP 125  
 |||||  
 RESULT 5  
 ADP19412  
 ID ADP19412 standard; protein; 163 AA.  
 XX  
 AC ADP19412;  
 XX  
 DT 26-AUG-2004 (first entry)  
 XX  
 DE Human secreted polypeptide #263.  
 XX  
 KW Human; secreted protein; genetic disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004110939-A1.  
 XX  
 PD 10-JUN-2004.  
 XX  
 PF 15-OCT-2001; 2001US-00978360.  
 XX  
 PR 17-DEC-1998; 98WO-IB002122.  
 PR 09-FEB-1999; 99WO-IB000282.  
 PR 21-JUN-2000; 2000WO-IB000951.  
 PR 15-SEP-2000; 2000US-00663600.  
 XX  
 PA (GEST ) GENSET SA.  
 XX  
 PI Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;





PT sarcoma.  
XX Example 2; SEQ ID NO 432; 210pp; English.  
XX  
CC The invention relates to a novel method for detecting soft tissue sarcoma  
CC which comprises obtaining a first soft tissue sample from an individual  
CC and a normal soft tissue sample from the same or different individual,  
CC determining the expression of a gene in both samples and comparing the  
CC expression of the gene in both soft tissue samples, where a higher level  
CC of protein expression in the first soft tissue sample indicates the  
CC presence of soft tissue sarcoma. The method of the invention has  
CC cytosstatic applications and may be useful for detecting soft tissue  
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
CC acid sequences may be useful in diagnostic and screening applications.  
CC The current sequence is that of a human soft tissue sarcoma-upregulated  
CC protein of the invention. The current sequence is not shown within the  
CC specification per se but was submitted in CD format by the inventor.  
XX  
SQ Sequence 553 AA;  
Query Match 30.5%; Score 67; DB 8; Length 553;  
Best Local Similarity 46.4%; Pred. No. 1.7;  
Matches 13; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
Qy 9 EDGWPASPTGLDPLTVRRHVPVAVVL 36  
Db 57 EEPWVPSGTDITLSTRTYRRNPGWSL 84  
RESULT 8  
ID ADN99926 standard; protein; 558 AA.  
XX  
AC ADN99926;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Novel human protein sequence #742.  
XX  
KW anti-inflammatory; dermatological; neuroprotective; immunomodulator;  
KW antibacterial; virucide; antipsoriatic; cytosstatic; gene therapy;  
KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;  
KW early aging; hormonal imbalance; ischemic heart disease;  
KW ulcerative colitis.  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO2004038003-A2.  
XX  
XX  
PD 06-MAY-2004.  
XX  
XX  
PF 24-OCT-2003; 2003WO-US033947.  
XX  
XX 25-OCT-2002; 2002US-0421061P.  
PR 25-OCT-2002; 2002US-0421080P.  
PR 25-OCT-2002; 2002US-0421552P.  
PR 25-OCT-2002; 2002US-0421614P.  
PR 30-OCT-2002; 2002US-0422177P.  
PR 30-OCT-2002; 2002US-0422178P.  
PR 15-NOV-2002; 2002US-0426355P.  
PR 15-NOV-2002; 2002US-0426384P.  
PR 15-NOV-2002; 2002US-0426394P.  
PR 15-NOV-2002; 2002US-0426430P.  
PR 15-NOV-2002; 2002US-0426916P.  
PR 27-NOV-2002; 2002US-0429224P.  
PR 27-NOV-2002; 2002US-0429275P.  
PR 27-NOV-2002; 2002US-0429302P.  
PR 27-NOV-2002; 2002US-0429326P.  
PR 04-DEC-2002; 2002US-0429651P.  
PR 04-DEC-2002; 2002US-0430651P.  
PR 04-DEC-2002; 2002US-0430657P.  
PR 04-DEC-2002; 2002US-0430663P.

PR 04-DEC-2002; 2002US-0430668P.  
PR 04-DEC-2002; 2002US-0430684P.  
PR 05-DEC-2002; 2002US-0430937P.  
PR 05-DEC-2002; 2002US-0430965P.  
PR 05-DEC-2002; 2002US-0431458P.  
PR 12-DEC-2002; 2002US-0433251P.  
PR 12-DEC-2002; 2002US-0433500P.  
PR 13-DEC-2002; 2002US-0433316P.  
PR 13-DEC-2002; 2002US-0433318P.  
PR 23-DEC-2002; 2002US-0436236P.  
PR 03-JAN-2003; 2003US-0437914P.  
PR 17-JAN-2003; 2003US-0440820P.  
PR 17-JAN-2003; 2003US-0440821P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476621P.  
PR 09-JUN-2003; 2003US-0476632P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485217P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 08-JUL-2003; 2003US-0485359P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;  
PI Wong JGP, Wu G, Zhang H, Zeng C;  
XX  
XX WPI; 2004-365511/34.  
DR N-PSDB; ADN99142.  
XX  
XX New nucleic acid molecules, useful in preparing a composition for  
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral  
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or  
PT ulcerative colitis.  
XX  
XX Claim 14; SEQ ID NO 1526; 532pp; English.  
PS  
XX The invention relates to a nucleic acid molecule comprising a  
CC polynucleotide sequence or its complement that encodes a polypeptide. The  
CC nucleic acid is useful in preparing a composition for treating or  
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,  
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic  
CC heart disease or ulcerative colitis. This sequence corresponds to a  
CC protein of the invention.  
XX  
SQ Sequence 558 AA;  
Query Match 30.5%; Score 67; DB 8; Length 558;  
Best Local Similarity 46.4%; Pred. No. 1.7;  
Matches 13; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 9 EDWVPASPDGTGLDPLTVRRHVPVAVWL 36  
 Db 62 EEPVPSGTDTLRTTYRRNPGSWL 89

RESULT 9  
 ABB64410  
 ID ABB64410 standard; protein; 205 AA.  
 AC ABB64410;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 20022.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN W0200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL08513.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 2022; 21pp + Sequence Listing; English.  
 CC  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 205 AA;  
 Query Match 29.1%; Score 64; DB 4; Length 205;  
 Best Local Similarity 30.0%; Pred. No. 1-3;  
 Matches 12; Conservative 10; Mismatches 16; Indels 2; Gaps 1;

QY 1 TKLRFLOTBDSWVPASPDGTGLDPLTVRRHVPVAVWL 40  
 Db 90 TRMAYTSVEHWKQW--EGLEBIIIVTRQIPTLHILMSLD 127

RESULT 10  
 ABF05282  
 ID ABF05282 standard; protein; 96 AA.  
 XX  
 AC ABP05282;  
 XX  
 DT 24-JUN-2002 (first entry)  
 XX

DE Human ORFX protein sequence SEQ ID NO:10546.  
 XX  
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KW hypertension; hypothyroidism; cholesterol ester storage disease;  
 KW immune deficiency; immune disorder; infectious disease;  
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KW myasthenia gravis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200192523-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 29-MAY-2001; 2001WO-US010836.  
 XX  
 PR 30-MAY-2000; 2000US-0206132P.  
 PR 29-AUG-2000; 2000US-0228716P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Leach MD;  
 XX  
 DR WPI; 2002-106308/14.  
 DR N-PSDB; ABN21034.  
 XX  
 PT Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders.  
 XX  
 PS Disclosure; SEQ ID NO 10546; 1037pp; English.  
 XX  
 CC The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
 CC treating or preventing a pathology associated with an ORFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
 CC sequences can be used in gene therapy. ORFX sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis.  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 96 AA;  
 Query Match 28.0%; Score 61.5; DB 5; Length 96;  
 Best Local Similarity 26.7%; Pred. No. 1.2;  
 Matches 16; Conservative 8; Mismatches 9; Indels 27; Gaps 3;

QY 1 TKLRFLO-----TEDS-----VPASPDGTGLDPLTVRRHVPVAVWL 36  
 Db 36 SRLRFVEQIRSVSPDSTTPSLGEVYTADSWGTLPLMLPGSADELS---HLFATWIL 92

RESULT 11  
 ABO64395



PR 07-JUN-2002; 2002JP-00167239.  
 XX (SUMO ) SUMITOMO CHEM CO LTD.  
 XX  
 XX Nakajima H, Mukumoto F, Takaishi M;  
 XX WPI; 2003-523102/49.  
 DR N-PSDB; ADC36139.  
 XX  
 PT Weed controller metabolism proteins deactivating porphyrinogen oxidase  
 PT (PPO)-inhibiting herbicides by N-demethylation and their genes, useful  
 PT e.g. in constructing new breeds of herbicide-resistant plants.  
 XX  
 XX Claim 1; SEQ ID NO 1; 813pp; Japanese.  
 XX  
 XX The invention relates to a novel DNA encoding a weed controller  
 CC metabolism protein. A protein of the invention has herbicide activity.  
 CC The proteins and their encoded genes are useful e.g. in constructing new  
 CC breeds of herbicide-resistant plants and also in developing various  
 CC agrochemicals. The present sequence is used in the exemplification of the  
 CC invention.  
 XX  
 XX SQ Sequence 408 AA;  
 Query Match 25.9%; Score 57; DB 7; Length 408;  
 Best Local Similarity 32.5%; Pred. No. 30;  
 Matches 13; Conservative 5; Mismatches 14; Indels 8; Gaps 1;  
 QY 9 EDSWVSPDPTGLDPLTVRRHP-----AVWVLLSRD 40  
 DB 21 QDRTCPFPPTGYDPLREARPLARVTLTYDGRAIWLVTGRD 60  
 RESULT 14  
 ADOL3847  
 ID ADOL3847 standard; protein; 4083 AA.  
 XX  
 XX ADOL3847;  
 DT 15-JUL-2004 (first entry)  
 DE Protein encoded by dbv gene cluster ORF17.  
 XX  
 KW dbv; gene cluster; Actinomadura sp.; biosynthesis; glycopeptide; A40926;  
 KW 4-hydroxy-phenylglycine; 3; 5-dihydroxy-phenylglycine;  
 KW heptapeptide skeleton; chlorination; aromatic residue;  
 KW beta-hydroxylation; tyrosine; cross-linking; N-acetyl glucuroamine;  
 KW mannosyl; export; antibiotic.  
 XX  
 OS Actinomadura sp. ATCC 39727.  
 XX  
 XX EP1413626-A1.  
 XX  
 PD 28-APR-2004.  
 XX  
 XX 23-OCT-2002; 2002EP-00023597.  
 XX  
 XX 23-OCT-2002; 2002EP-00023597.  
 XX  
 XX (VICU-) VICURON PHARM INC.  
 XX  
 XX Donadio S, Sosio M, Beltrametti F;  
 XX WPI; 2004-332813/31.  
 DR N-PSDB; ADOL3830.  
 XX  
 XX Novel isolated polypeptide A40926, comprising polypeptide sequence or  
 PT open reading frame polypeptide involved in biosynthetic pathway of  
 PT A40926, useful for synthesizing glycopeptide antibiotic A40926.  
 XX  
 XX Claim 28; Page 95-112; 164pp; English.  
 XX  
 XX This sequence is encoded by the dbv gene cluster from Nonomura sp.

CC ATCC39727 (formerly Actinomadura sp. ATCC39727). The 37 proteins encoded  
 CC by the gene cluster are involved in the biosynthesis of the glycopeptide  
 CC A40926. dbv ORF 1, 2, 5, and 37 encode polypeptides required for the  
 CC synthesis of 4-hydroxy-phenylglycine residues of A40926. dbv ORF 30-34  
 CC and 37 encode polypeptides required for the synthesis of 3, 5-dihydroxy-  
 CC phenylglycine residues of A40926. dbv ORF 16, 17, 25, 26 and 36 encode  
 CC polypeptides required for the synthesis of heptapeptide skeleton of  
 CC A40926. dbv ORF 10 encodes polypeptide required for the chlorination of  
 CC the aromatic residues of A40926. dbv ORF 28 encodes a polypeptide  
 CC required for the beta-hydroxylation of the tyrosine residue of amino acid  
 CC 6 of A40926. dbv ORF 11-14 encode polypeptides required for cross-linking  
 CC of the aromatic residues of amino acids at positions 2 and 4, 4 and 6, 1  
 CC and 3, and 5 and 7 of A40926. dbv ORF 9, 23 and 29 encode polypeptides  
 CC required for addition and formation of N-acetyl glucuroamine residue of  
 CC A40926. dbv ORF 20 or 27 encode polypeptides required for the attachment  
 CC of mannosyl residues or N-methylation of A40926. dbv ORF 7, 18, 19, 24  
 CC and 35 encode polypeptides required for the export of A40926 or its  
 CC precursors. dbv ORF 3, 4, 6, and 22 encode polypeptides required for  
 CC regulating the expression of one or more genes of the dbv gene cluster.  
 CC The dbv gene cluster and the proteins encoded by it are useful for  
 CC producing glycopeptide antibiotic A40926 or its precursor.  
 XX  
 XX SQ Sequence 4083 AA;  
 Query Match 25.7%; Score 56.5; DB 8; Length 4083;  
 Best Local Similarity 40.0%; Pred. No. 5.7e+02;  
 Matches 16; Conservative 3; Mismatches 12; Indels 9; Gaps 1;  
 QY 11 SNWPASPTGLDPLTVRRH-----VPAWVLLSRDP 41  
 DB 2385 AYVTAEGSGLDPAVREHLAAGRLPEFMVPAWVLLDGV 2424  
 RESULT 15  
 ABG21326  
 ID ABG21326 standard; protein; 264 AA.  
 XX  
 XX ABG21326;  
 DT 18-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #21317.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 XX 30-MAR-2001; 2001WO-US008631.  
 XX  
 XX 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS85513.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 XX Claim 20; SEQ ID NO 51685; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain

DS

Sequence 264 AA;

Query Match 25.5%; Score 56; DB 4; Length 264;

Best Local Similarity 38.7%; Pred. No. 25;  
Matches 12; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

16 SPDTGLDPLTVRRHVP-----AVVLLSRD 40  
Qy

197 APQTLINPMEVREHVPRRHASQCFFWVKSGRD 227

Search completed: January 26, 2005, 15:01:07  
Job time : 42.861 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:06:38 ; Search time 34.1898 Seconds  
(without alignments)  
433.254 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_85\_125

Perfect score: 220

Sequence: 1 TKLRFQTESWVPASPDTC.....DPLTVRRHVPVAVWLLSRDP 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220	100.0	163	9	US-09-731-872-459
2	220	100.0	163	10	US-09-876-997-459
3	194	88.2	163	11	US-09-978-360A-668
4	67	30.5	553	17	US-10-723-860-432
5	65	29.5	242	17	US-10-425-115-254668
6	64	29.1	276	17	US-10-425-115-308895
7	64	29.1	276	17	US-10-425-115-308929
8	64	29.1	278	15	US-10-425-114-60670
9	64	29.1	301	15	US-10-425-114-61107
10	64	29.1	315	15	US-10-425-114-54315
11	64	29.1	315	15	US-10-425-114-58679
12	64	29.1	325	15	US-10-425-114-60913
13	63.5	28.9	436	14	US-10-156-761-11365

14	62	28.2	261	17	US-10-425-115-308917	Sequence 308917,
15	62	28.2	275	16	US-10-767-701-45240	Sequence 45240, A
16	60.5	27.5	539	17	US-10-425-115-312091	Sequence 312091,
17	58	26.4	57	16	US-10-437-963-120370	Sequence 120370,
18	58	26.4	1037	16	US-10-408-765A-2808	Sequence 2808, Ap
19	57	25.9	75	17	US-10-425-115-234349	Sequence 234349,
20	57	25.9	99	17	US-10-425-115-303949	Sequence 303949,
21	57	25.9	262	16	US-10-437-963-191078	Sequence 191078,
22	57	25.9	298	15	US-10-425-114-58193	Sequence 58193, A
23	56.5	25.7	201	17	US-10-425-115-319295	Sequence 319295,
24	56.5	25.7	257	17	US-10-425-115-213299	Sequence 213299,
25	56.5	25.7	444	16	US-10-437-963-121940	Sequence 121940,
26	56.5	25.7	788	14	US-10-369-493-11652	Sequence 11652, A
27	56.5	25.7	788	14	US-10-369-493-11652	Sequence 11652, A
28	56.5	25.7	1446	15	US-10-087-684-44	Sequence 14699, A
29	56.5	25.7	1446	15	US-10-087-684-44	Sequence 14699, A
30	56	25.5	245	15	US-10-218-779-44	Sequence 44, Appl
31	56	25.5	245	15	US-10-218-779-44	Sequence 44, Appl
32	56	25.5	263	16	US-10-437-963-152741	Sequence 229595,
33	56	25.5	281	15	US-10-424-599-237821	Sequence 152741,
34	56	25.5	287	16	US-10-437-963-152743	Sequence 237821,
35	55.5	25.2	804	14	US-10-369-493-13782	Sequence 152743,
36	55.5	25.2	324	14	US-10-156-761-13399	Sequence 12782, A
37	55.5	25.2	402	15	US-10-282-122A-52780	Sequence 13399, A
38	55.5	25.2	619	15	US-10-087-684-42	Sequence 52780, A
39	55.5	25.2	1327	16	US-10-755-889-510	Sequence 42, Appl
40	55.5	25.2	1446	15	US-10-087-684-12	Sequence 510, Appl
41	55.5	25.2	1446	15	US-10-218-779-12	Sequence 12, Appl
42	55.5	25.2	1451	17	US-10-483-512-22	Sequence 22, Appl
43	55.5	25.2	2341	15	US-10-087-684-43	Sequence 43, Appl
44	55.5	25.2	2341	15	US-10-218-779-43	Sequence 43, Appl
45	55	25.0	888	16	US-10-437-963-197785	Sequence 197785,

#### ALIGNMENTS

##### RESULT 1

US-09-731-872-459  
; Sequence 459, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 459  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-731-872-459

Query Match 100.0%; Score 220; DB 9; Length 163;  
Best Local Similarity 100.0%; Pred No. 6.2e-21;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKLRFQTESWVPASPDTCGLDPLTVRRHVPVAVWLLSRDP 41

Db 85 TKLRFQTESWVPASPDTCGLDPLTVRRHVPVAVWLLSRDP 125

##### RESULT 2

US-09-876-997-459  
; Sequence 459, Application US/09876997

```
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.USA.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 459
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-876-997-459

Query Match      100.0%; Score 220; DB 10; Length 163;
Best Local Similarity 100.0%; Pred. No. 6.2e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLRFQTEDSWVPASPDGDLPLTVRRHVPVAVWLLSRDP 41
   |||||
Db 85 TKLRFQTEDSWVPASPDGDLPLTVRRHVPVAVWLLSRDP 125

RESULT 3
US-09-978-360A-668
; Sequence 668, Application US/09/978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.USA.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 668
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -52...-1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36, 47, 54, 65, 70, 103, 107, 108)
; OTHER INFORMATION: unknown
US-09-978-360A-668

Query Match      88.2%; Score 194; DB 11; Length 163;
Best Local Similarity 87.8%; Pred. No. 1.7e-17;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TKLRFQTEDSWVPASPDGDLPLTVRRHVPVAVWLLSRDP 41
   |||||
Db 85 TKLRFQTEDSWVPASPDGDLPLTVRRHVPVAVWLLSRDP 125

RESULT 4
US-10-723-860-432
; Sequence 432, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginzburg, Wendy M.
; APPLICANT: Zlocznik, Albert
; TITLE OF INVENTION: Methods of diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 432
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-432

Query Match      30.5%; Score 67; DB 17; Length 553;
Best Local Similarity 46.4%; Pred. No. 3.9;
Matches 13; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 9 EDSWVPASPDGDLPLTVRRHVPVAVWL 36
   |||
Db 57 BEPWVPGTDTTSLRTTYRRNPGWSL 84

RESULT 5
US-10-425-115-254668
; Sequence 254668, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 254668
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(242)
; OTHER INFORMATION: unsure at all Xaa locations
```



; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_163838C.1.pep  
US-10-425-115-254668

Query Match 29.1%; Score 65; DB 17; Length 242;  
Best Local Similarity 34.1%; Pred. No. 2.8; Mismatches 17; Indels 2; Gaps 1;  
Matches 14; Conservative 8;

QY 1 TKLRFLOEDSWPASPDTGLDPLTVRRHVPVAVVLLSRDP 41

DB 78 TSIESVGITDWEPL--EGLVPLETTRHVSMTITLSKEP 116

## RESULT 6

US-10-425-115-308895  
; Sequence 308895, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 308895  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_4477C.1.pep  
US-10-425-115-308895

Query Match 29.1%; Score 64; DB 17; Length 276;  
Best Local Similarity 43.8%; Pred. No. 4.4;  
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 10 DSWVPASPDTGLDPLTVRRHVPVAVVLLSRDP 41

DB 87 DTWEPL--EGLPLETTRHVSMTITLSKNP 116

## RESULT 7

US-10-425-115-308929  
; Sequence 308929, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 308929  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_4480C.1.pep  
US-10-425-115-308929

Query Match 29.1%; Score 64; DB 17; Length 276;  
Best Local Similarity 43.8%; Pred. No. 4.4;  
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 10 DSWVPASPDTGLDPLTVRRHVPVAVVLLSRDP 41

DB 87 DTWEPL--EGLPLETTRHVSMTITLSKNP 116

## RESULT 8

US-10-425-114-60670  
; Sequence 60670, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 60670  
; LENGTH: 278  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3357-033-B9\_FLI.pep  
US-10-425-114-60670

Query Match 29.1%; Score 64; DB 15; Length 278;  
Best Local Similarity 43.8%; Pred. No. 4.4;  
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 10 DSWVPASPDTGLDPLTVRRHVPVAVVLLSRDP 41

DB 89 DTWEPL--EGLPLETTRHVSMTITLSKNP 118

## RESULT 9

US-10-425-114-61107  
; Sequence 61107, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 61107  
; LENGTH: 301  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3332-020-F9\_FLI.pep  
US-10-425-114-61107

Query Match 29.1%; Score 64; DB 15; Length 301;  
Best Local Similarity 43.8%; Pred. No. 4.8;  
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 10 DSWVPASPDTGLDPLTVRRHVPVAVVLLSRDP 41

DB 112 DTWEPL--EGLPLETTRHVSMTITLSKNP 141

## RESULT 10

```

US-10-425-114-54515
; Sequence 54515, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54515
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17184G05_FLI.pep
US-10-425-114-54515

Query Match      29.1%; Score 64; DB 15; Length 315;
Best Local Similarity 43.8%; Pred. No. 5.1;
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 10 DSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 126 DTWEPL--EGLLPLETRRHVSMITVTLKNP 155

RESULT 11
US-10-425-114-58679
; Sequence 58679, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58679
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700577207_FLI.pep
US-10-425-114-58679

Query Match      29.1%; Score 64; DB 15; Length 315;
Best Local Similarity 43.8%; Pred. No. 5.1;
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 10 DSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 126 DTWEPL--EGLLPLETRRHVSMITVTLKNP 155

RESULT 12
US-10-425-114-60913
; Sequence 60913, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

```

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60913
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-105-A5_FLI.pep
US-10-425-114-60913

Query Match      29.1%; Score 64; DB 15; Length 325;
Best Local Similarity 43.8%; Pred. No. 5.3;
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 10 DSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 136 DTWEPL--EGLLPLETRRHVSMITVTLKNP 165

RESULT 13
US-10-156-761-11365
; Sequence 11365, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11365
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11365

Query Match      28.9%; Score 63.5; DB 14; Length 436;
Best Local Similarity 41.0%; Pred. No. 8.5;
Matches 16; Conservative 4; Mismatches 16; Indels 3; Gaps 1;

QY 3 LRFQTDSDWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 172 LSWIVTEHWRPAVTVSLAALAV---VPPFWLLLRDHP 207

RESULT 14
US-10-425-115-308917
; Sequence 308917, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

```

Search completed: January 26, 2005, 15:28:37  
Job time : 35.1898 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:47 ; Search time 10.4237 Seconds  
(without alignments)  
260.851 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_85\_125

Perfect score: 220  
Sequence: 1 TKLRFLOTDSWVPASPDTC.....DPLTVRRHVPVWVLLSRDP 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	96.8	128	4	US-09-621-976-4077
2	61.5	28.0	586	4	US-09-489-039A-10912
3	54	24.5	1282	4	US-09-489-039A-9644
4	53	24.1	409	4	US-09-252-991A-29674
5	51.5	23.4	400	4	US-09-252-991A-31900
6	51.5	23.4	452	4	US-09-489-039A-12558
7	51.5	23.4	497	3	US-09-377-557-10
8	51.5	23.4	602	2	US-08-419-652-6
9	51.5	23.4	717	4	US-09-489-039A-13855
10	51.5	23.4	771	1	US-07-923-976-6
11	51.5	23.4	783	6	5422248-2
12	51.5	23.4	836	1	US-07-923-976-4
13	51.5	23.4	863	1	US-07-923-976-8
14	51	23.2	196	3	US-09-134-001C-4956
15	50.5	23.0	837	1	US-07-923-976-2
16	50	22.7	334	4	US-09-270-767-35677
17	50	22.7	334	4	US-09-270-767-50894
18	50	22.7	439	4	US-09-248-796A-15968
19	50	22.7	585	4	US-09-489-039A-14060
20	50	22.7	817	2	US-08-820-170A-28
21	50	22.7	817	3	US-09-055-699-28
22	50	22.7	817	3	US-09-273-565-28
23	50	22.7	817	3	US-09-565-538-28
24	50	22.7	817	3	US-09-661-468-28
25	50	22.7	817	4	US-09-976-165-28
26	50	22.7	829	2	US-08-820-170A-31
27	50	22.7	829	3	US-09-055-699-31

28	50	22.7	829	3	US-09-273-565-31	Sequence 31, Appl
29	50	22.7	829	3	US-09-565-538-31	Sequence 31, Appl
30	50	22.7	829	3	US-09-661-468-31	Sequence 31, Appl
31	50	22.7	829	4	US-09-976-165-31	Sequence 31, Appl
32	49.5	22.5	539	4	US-09-252-991A-26939	Sequence 26939, A
33	49.5	22.5	688	4	US-09-489-039A-9813	Sequence 9813, Ap
34	49.5	22.5	942	2	US-08-884-681-4	Sequence 4, Appl
35	49.5	22.5	942	3	US-08-560-005-6	Sequence 6, Appl
36	49.5	22.5	942	3	US-09-258-643-4	Sequence 4, Appl
37	49.5	22.5	942	3	US-09-418-540-6	Sequence 6, Appl
38	49.5	22.5	942	4	US-09-969-528-6	Sequence 6, Appl
39	49	22.3	87	4	US-09-621-976-6870	Sequence 6870, Ap
40	49	22.3	373	4	US-09-149-476-374	Sequence 374, App
41	49	22.3	404	4	US-09-266-965-103	Sequence 103, App
42	49	22.3	457	4	US-09-252-991A-20955	Sequence 20955, A
43	49	22.3	602	4	US-09-489-039A-12436	Sequence 12436, A
44	49	22.3	1019	4	US-09-434-066-23	Sequence 23, Appl
45	49	22.3	2353	3	US-08-984-709A-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-4077  
; Sequence 4077, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 4077  
; LENGTH: 128  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -52..-1  
; NAME/KEY: UNSURE  
; LOCATION: 72  
; OTHER INFORMATION: Xaa = Asp,Glu,Gly,Val  
US-09-621-976-4077

Query Match 96.8%; Score 213; DB 4; Length 128;  
Best Local Similarity 97.6%; Pred. No. 2.5e-22;  
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TKLRFLOTDSWVPASPDTCGLDPLTVRRHVPVWVLLSRDP 41  
|||  
Db 85 TKLRFLOTDSWVPASPDTCGLDPLTVRRHVPVWVLLSRXP 125

RESULT 2

US-09-489-039A-10912  
; Sequence 10912, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLSBIELLA  
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10912





REFERENCE/DOCKET NUMBER: 514853  
TELEPHONE: 703-415-1500  
TELEFAX: 703-415-1508  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 771 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-923-976-6

Query Match 23.4%; Score 51.5; DB 1; Length 771;  
Best Local Similarity 43.5%; Pred. No. 1e+02;  
Matches 10; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 12 WVPASPDGTGLDPLTVRRHVPVW 34  
||| | : | ||| |  
Db 547 WVPPELKGKSLT---HYTIFW 566

RESULT 11  
5422248-2  
Patent No. 5422248  
APPLICANT: SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;  
BENSON, CURTIS M.  
TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY  
STIMULATING FACTOR RECEPTORS  
NUMBER OF SEQUENCES: 6  
CURRENT APPLICATION DATA:  
FILING DATE: 15-JAN-1993  
PRIOR APPLICATION NUMBER: 587,329  
FILING DATE: 24-SEP-1990  
APPLICATION NUMBER: 522,952  
FILING DATE: 03-APR-1990  
APPLICATION NUMBER: 416,306  
FILING DATE: 03-OCT-1989  
APPLICATION NUMBER: 412,816  
FILING DATE: 26-SEP-1989  
SEQ ID NO: 2:  
LENGTH: 783  
5422248-2

Query Match 23.4%; Score 51.5; DB 6; Length 783;  
Best Local Similarity 43.5%; Pred. No. 1e+02;  
Matches 10; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 12 WVPASPDGTGLDPLTVRRHVPVW 34  
||| | : | ||| |  
Db 547 WVPPELKGKSLT---HYTIFW 566

RESULT 12  
US-07-923-976-4  
Sequence 4, Application US/07923976  
Patent No. 5574136  
GENERAL INFORMATION:  
APPLICANT: Nagata, Shigekazu  
APPLICANT: Fukunaga, Rikio  
TITLE OF INVENTION: DNA Encoding Granulocyte  
Stimulating Factor Receptor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones, Tullar & Cooper, P.C.  
STREET: P.O. Box 2266 Eads Station  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/923,976  
FILING DATE: 19920922  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 74539/1990  
FILING DATE: 23-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 176629/1990  
FILING DATE: 03-JUL-1990  
APPLICATION DATA:  
PCT/JP91/00375  
FILING DATE: 22-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hellwege, James W.  
REGISTRATION NUMBER: 28,808  
REFERENCE/DOCKET NUMBER: 514853  
TELEPHONE: 703-415-1500  
TELEFAX: 703-415-1508  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 836 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-923-976-4

Query Match 23.4%; Score 51.5; DB 1; Length 836;  
Best Local Similarity 43.5%; Pred. No. 1e+02;  
Matches 10; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 12 WVPASPDGTGLDPLTVRRHVPVW 34  
||| | : | ||| |  
Db 547 WVPPELKGKSLT---HYTIFW 566

RESULT 13  
US-07-923-976-8  
Sequence 8, Application US/07923976  
Patent No. 5574136  
GENERAL INFORMATION:  
APPLICANT: Nagata, Shigekazu  
APPLICANT: Fukunaga, Rikio  
TITLE OF INVENTION: DNA Encoding Granulocyte  
Stimulating Factor Receptor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones, Tullar & Cooper, P.C.  
STREET: P.O. Box 2266 Eads Station  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/923,976  
FILING DATE: 19920922  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 74539/1990  
FILING DATE: 23-MAR-1990  
APPLICATION DATA:  
PCT/JP91/00375  
FILING DATE: 22-MAR-1991



```

; MEDIUM TYPE: Floppy disk

```

Search completed: January 26, 2005, 15:08:53  
Job time : 11.4237 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:44:05 ; Search time 39 Seconds  
(without alignments)  
101.151 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_85\_125  
Perfect score: 220  
Sequence: 1 TKLRFLOTEDSWVPSPDTG.....DPLTVRRHVPVWVLLSRDP 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 9448

Minimum DB seq length: 0  
Maximum DB seq length: 41

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	18.2	36	2 T36994	probable transposase
2	37	16.8	35	2 H84214	hypothetical prote
3	36	16.4	34	2 S46474	lg heavy chain v r
4	36	16.4	36	1 LBRFGV	light-harvesting p
5	35.5	16.1	38	2 G45095	photosystem I lig
6	35	15.9	25	2 S03275	photosystem II 22K
7	34	15.5	35	2 A61077	glue protein - rib
8	34	15.5	40	2 F45095	photosystem I lig
9	33.5	15.2	30	2 E45095	photosystem I lig
10	33.5	15.2	31	2 I61598	myosin - human (fr
11	33	15.0	18	2 S14661	photosystem I prot
12	33	15.0	24	2 A61505	pollen allergen Ph
13	32.5	14.8	38	2 I70022	kallikrein - mouse
14	32	14.5	20	2 S68341	procathepsin L - g
15	32	14.5	28	2 S20393	trypsin inhibitor
16	32	14.5	35	2 D82224	hypothetical prote
17	31.5	14.3	35	2 S46472	lg heavy chain v r
18	31	14.1	20	2 JN0252	water-soluble 35K
19	31	14.1	31	2 T29628	hypothetical prote
20	31	14.1	34	2 S46469	lg heavy chain v r
21	31	14.1	39	2 S26937	lg heavy chain v r
22	31	14.1	40	2 G69471	hypothetical prote
23	30.5	13.9	38	2 I70018	kallikrein - mouse
24	30	13.6	20	2 PQ0071	T-cell receptor be
25	30	13.6	30	2 S55462	mers protein homol
26	30	13.6	35	2 S26887	lg heavy chain v r
27	30	13.6	35	2 S46473	lg heavy chain v r
28	30	13.6	35	2 C95037	hypothetical prote
29	30	13.6	35	2 S58708	neutral phosphatas

hypothetical prote  
histone H3 homolog  
hypothetical prote  
hypothetical prote  
hypothetical prote  
thyrotropin-releas  
insulin-like growt  
apolipoprotein Cbl  
agglutinin - Japan  
serine proteinase  
trypsin inhibitor  
dopamine D3 recept  
hypothetical prote  
cytochrome c (EDH)  
hypothetical EWSR1  
homeobox 4 protein  
antifreeze glycopr

ALIGNMENTS

RESULT 1

T36994  
probable transposase, truncated section 1 [imported] - Streptomyces coelicolor (fragmen  
C;Species: Streptomyces coelicolor  
C;Date: 08-Sep-2000 #sequence\_revision 08-Sep-2000 #text\_change 15-Sep-2000  
C;Accession: T36994

R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1999

A;Reference number: Z21618

A;Accession: T36994

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-36 <OJJ>

A;Cross-references: EMBL:AL109949; PIDN:CAB52908.1; GSPDB:GN00070; SCOEDB:SCJ11.23c

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SCJ11.23c

C;Superfamily: Streptomyces coelicolor probable transposase SC609.06c

Query Match 18.2%; Score 40; DB 2; Length 36;

Best Local Similarity 40.6%; Pred. No. 1.2e+02;

Matches 13; Conservative 3; Mismatches 8; Indels 8; Gaps 2;

QY 8 TEDSWVPSPDTGLDP-----LTVRRHVPV 33

Db :|||:|||:|||:|||

2 SEPSW--TAPFTGLSPRCWKLVTLRQGV 31

RESULT 2

H84214  
hypothetical protein Vng0570h [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C;Accession: H84214

R;Ng, W.V.; Kennedy, S.P.; Mahaitas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,

; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L

A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: H84214

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-35 <STO>

A;Cross-references: UNIPROT:Q9HRS4; GB:AE004437; NID:g10580164; PIDN:AAG19084.1; GSPDB

C;Genetics:

A;Gene: VNG0570H

Query Match 16.8%; Score 37; DB 2; Length 35;

Best Local Similarity 46.7%; Pred. No. 2.9e+02;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 15 ASPDTGLDPLTVRRH 29  
: : : : :  
Db 15 AAADQGIDPWTDH 29

## RESULT 3

S46474  
Ig heavy chain V region (DA-9) - human  
C:Species: Homo sapiens (man)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Jun-2000  
C:Accession: S46474  
R:Cook, G.P.; Tomlinson, I.M.; Walter, G.; Riethman, H.; Carter, N.P.; Buluwela, L.; Wirtz, N.; et al.  
A:Title: A map of the human immunoglobulin V(H) locus completed by analysis of the telomeric region  
A:Reference number: S46460; MUID:95004581; PMID:7920635  
A:Accession: S46474  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-34 <COO>  
A:CROSS-references: EMBL:Z29984; NID:G505470; PIDN:CAA82874.1; PID:G1335184  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 16.4%; Score 36; DB 2; Length 34;  
Best Local Similarity 43.8%; Pred. No. 3.8e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 11 SWVPASPDTGLDPLTV 26  
: : : : :  
Db 8 SWVRQAPGKLEWVS 23

## RESULT 4

LBRRGV  
light-harvesting protein B-1015 gamma chain - Rhodospseudomonas viridis  
C:Species: Rhodospseudomonas viridis  
C>Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 09-Jul-2004  
C:Accession: A03455  
R:Brüscholz, R.A.; Jay, F.; Suter, F.; Zuber, H.  
A:Title: Chem. Hoppe-Seyler 366, 87-98, 1985  
A:Reference number: A90694; MUID:85225948; PMID:3890891  
A:Accession: A03455  
A:Molecule type: protein  
A:Residues: 1-36 <BR>  
A:CROSS-references: UNIPROT:P04126  
A:Note: residue 34 is Val in approximately one-third of the molecules  
C:Superfamily: light-harvesting protein gamma chain  
C:Keywords: antenna complex; bacteriochlorophyll; membrane protein

Query Match 16.4%; Score 36; DB 1; Length 36;  
Best Local Similarity 28.6%; Pred. No. 4.1e+02;  
Matches 10; Conservative 8; Mismatches 15; Indels 2; Gaps 2;

QY 5 FLOTEDSWVPASPDGL-DPLTVRRHVPVWVLLS 38  
: : : : :  
Db 1 YPAADGSVVPISDWNLVPLGI-LGIPTIWIALT 34

## RESULT 5

G45095  
photosystem I light-harvesting complex chlorophyll a/b protein p22.1/p22 - Chlamydomonas reinhardtii  
C:Species: Chlamydomonas reinhardtii  
C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: G45095  
R:Basai, R.; Soen, S.Y.; Frank, G.; Zuber, H.; Roehaix, J.D.  
A:Title: Chem. 267, 25714-25721, 1992  
A:Reference number: A45095; MUID:93100280; PMID:1464588  
A:Accession: G45095  
A>Status: preliminary  
A:Molecule type: protein

A:Residues: 1-38 <BAS>  
A:CROSS-references: UNIPROT:Q9S8T9  
A:Note: sequence extracted from NCBI backbone (NCBIP:120940)  
C:Superfamily: chlorophyll a/b-binding protein  
C:Keywords: chloroplast; thylakoid

Query Match 16.1%; Score 35.5; DB 2; Length 38;  
Best Local Similarity 33.3%; Pred. No. 5e+02;  
Matches 9; Conservative 4; Mismatches 3; Indels 11; Gaps 2;

QY 11 SWVPLS-----PDT-----GLDPLTV 26  
: : : : :  
Db 4 NWLPGSDAPAWLPDDLPGNYGDFPLSL 30

## RESULT 6

S03275  
photosystem II 22K protein - spinach (fragment)  
C:Species: Spinacia oleracea (spinach)  
C>Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 21-Aug-1998  
C:Accession: S03275  
R:Murata, N.; Kajimura, H.; Fujimura, Y.; Miyao, M.; Murata, T.; Watanabe, A.; Shinozaki, K.; et al.  
A:Title: Partial amino acid sequences of the proteins of pea and spinach photosystem II  
A:Reference number: S03269  
A:Accession: S03275  
A:Molecule type: protein  
A:Residues: 1-25 <MUR>  
C:Superfamily: chlorophyll a/b-binding protein  
C:Keywords: chloroplast; photosynthesis; photosystem II

Query Match 15.9%; Score 35; DB 2; Length 25;  
Best Local Similarity 53.3%; Pred. No. 3.7e+02;  
Matches 8; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 11 SWVPASPDTG--LDP 23  
: : : : :  
Db 8 SWIPAVKGGNPLDP 22

## RESULT 7

A61077  
glue protein - ribbed mussel (fragments)  
C:Species: Geukensia demissa (ribbed mussel)  
C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Jul-2004  
C:Accession: A61077  
R:Waite, J.H.; Hansen, D.C.; Little, K.T.  
A:Title: Comp. Physiol. B 159, 517-525, 1989  
A:Reference number: A61077; MUID:90110595; PMID:2481690  
A:Accession: A61077  
A:Molecule type: protein  
A:Residues: 1-35 <WAI>  
A:CROSS-references: UNIPROT:Q7M3X7  
C:Superfamily: glutenin  
C:Keywords: hydroxylysine; hydroxyproline; tandem repeat  
F:3,7,12,21/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental  
F:14/Modified site: 4-hydroxyproline (pro) #status experimental  
F:16,30,34/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) (partial) #status experimental

Query Match 15.5%; Score 34; DB 2; Length 35;  
Best Local Similarity 40.0%; Pred. No. 7.2e+02;  
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 9 EDSWVPASPDPTGLDP 23  
: : : : :  
Db 9 QTGVPGYKQTGYDP 23

## RESULT 8

F45095  
photosystem I light-harvesting complex chlorophyll a/b protein p18 - Chlamydomonas reinhardtii  
C:Species: Chlamydomonas reinhardtii

C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: F45095  
R;Bassi, R.; Soen, S.Y.; Frank, G.; Zuber, H.; Rochaix, J.D.  
J. Biol. Chem. 267, 25714-25721, 1992  
A;Title: Characterization of chlorophyll a/b proteins of photosystem I from Chlamydomonas reinhardtii  
A;Reference number: A45095; MUID:93100280; PMID:1464588  
A;Accession: F45095  
A>Status: preliminary  
A:Molecule type: protein  
A;Residues: 1-40 <BAS>  
A;Cross-references: UNIPROT:Q9S8UO  
A;Note: sequence extracted from NCBI backbone (NCBIP:120939)

Query Match 15.5%; Score 34; DB 2; Length 40;  
Best Local Similarity 28.3%; Pred. No. 8.3e+02;  
Matches 13; Conservative 2; Mismatches 5; Indels 26; Gaps 3;

QY 11 SWVPASP-----DT-----GLDPLTVRRHVPVWLLSRDP 41  
||| || | ||| | : ||  
DB 3 SWLPGSGI PAHLTPAAQAALAGNFGFDPLG-----LGKDP 37  
||| || | ||| | : ||

RESULT 9  
E45095  
photosystem I light-harvesting complex chlorophyll a/b protein p18.1 - Chlamydomonas reinhardtii  
C;Species: Chlamydomonas reinhardtii  
C;Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: E45095  
R;Bassi, R.; Soen, S.Y.; Frank, G.; Zuber, H.; Rochaix, J.D.  
J. Biol. Chem. 267, 25714-25721, 1992  
A;Title: Characterization of chlorophyll a/b proteins of photosystem I from Chlamydomonas reinhardtii  
A;Reference number: A45095; MUID:93100280; PMID:1464588  
A;Accession: E45095  
A>Status: preliminary  
A:Molecule type: protein  
A;Residues: 1-30 <BAS>  
A;Cross-references: UNIPROT:Q9S8UO  
A;Note: sequence extracted from NCBI backbone (NCBIP:120937)  
C;Superfamily: chlorophyll a/b-binding protein

Query Match 15.2%; Score 33.5; DB 2; Length 30;  
Best Local Similarity 33.3%; Pred. No. 7e+02;  
Matches 8; Conservative 1; Mismatches 4; Indels 11; Gaps 1;

QY 12 WVPASP-----DTGLDPL 24  
||| ||| ||| ||| |||  
DB 7 WLPGSTPXAHLKGDLPGDFGFDPL 30  
||| ||| ||| ||| |||

RESULT 10  
I61698  
myosin - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004  
C/Accession: I61698  
R;Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.  
Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994  
A;Title: Identification and overlapping expression of multiple unconventional myosin genes  
A;Reference number: A55758; MUID:94294418; PMID:8022818  
A;Accession: I61698  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A;Residues: 1-31 <RES>  
A;Cross-references: UNIPROT:Q14786; GB:I29147; NID:g457255; PIDN:AAA20910.1; PID:g531140

Query Match 15.2%; Score 33.5; DB 2; Length 31;  
Best Local Similarity 50.0%; Pred. No. 7.3e+02;  
Matches 7; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

QY 3 LRFLQT---EDSWV 13  
|:| |:| |:| |:| |:|  
DB 7 LQFLATISQGDSWI 20  
|:| |:| |:| |:| |:|

Query Match 14.8%; Score 32.5; DB 2; Length 38;  
Best Local Similarity 50.0%; Pred. No. 1.2e+03;  
Matches 8; Conservative 3; Mismatches 4; Indels 1; Gaps 1;  
  
QY 21 LDPLTVRR-HVPAVVV 35  
|: ||||| |:  
DB 11 LEDLTVRRIPNCGMWL 26  
  
RESULT 14  
S68341  
procathepsin L - guinea pig (fragments)  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 17-Mar-1999  
C:Accession: S68341  
R:McDonald, J.K.; Emerick, J.M.C.  
Arch. Biochem. Biophys. 323, 409-422, 1995  
A:Title: Purification and characterization of procathepsin L, a self-processing zymogen  
A:Reference number: S68341; MUID:96063618; PMID:7487106  
A:Accession: S68341  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-11;12-20 <MCD>

Query Match 14.5%; Score 32; DB 2; Length 20;  
Best Local Similarity 53.8%; Pred. No. 7.1e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 7 QTEDSNVPASPD 19  
|| |:  
DB 6 QTLDAQLPKSVDT 18  
  
RESULT 15  
S20393  
trypsin inhibitor A - balsam pear  
C:Species: Momordica charantia (balsam pear, bitter melon)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S20393  
R:Huang, Q.; Liu, S.; Tang, Y.; Zeng, F.; Qian, R.  
FEBS Lett. 297, 143-146, 1992  
A:Title: Amino acid sequencing of a trypsin inhibitor by refined 1.6 A X-ray crystal structure  
A:Reference number: S20393; MUID:92201369; PMID:1551419  
A:Accession: S20393  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-28 <HUA>  
A:Cross-references: UNIPROT:P30709  
C:Superfamily: squash trypsin inhibitor ITD I

Query Match 14.5%; Score 32; DB 2; Length 28;  
Best Local Similarity 38.5%; Pred. No. 1e+03;  
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
QY 28 RHVPAVWLLSRD 40  
|: |:  
DB 1 RSCPRIWMECTRD 13

Search completed: January 26, 2005, 15:53:55  
Job time : 40 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	47.5	21.6	37	2	Q73H28	Q73h28 wolbachia p
2	47.5	21.6	37	2	AAS14438	AAS14438 wolbachia
3	39	17.7	32	2	Q8NG01	Q8ng01 homo sapien
4	39	17.7	32	2	Q8BLG0	Q8blg0 pseudomonas
5	39	17.7	32	2	Q8VKN1	Q8vkn1 mycobacteri
6	39	17.7	39	2	Q6LDS8	Q6lds8 cryptotolagu
7	39	17.7	39	2	AAA31172	AAA31172 cryptotolag
8	38	17.3	24	2	Q9SCH7	Q9sch7 semiliquida
9	38	17.3	24	2	Q9SCH9	Q9sch9 atlingia gr
10	38	17.3	24	2	Q9SC11	Q9sc11 atlingia ta
11	37.5	17.0	26	2	Q6E554	Q6e554 hepatitis c
12	37	16.8	21	2	Q8SUR5	Q8sur5 strephonema
13	37	16.8	21	2	Q8SUR6	Q8sur6 strephonema
14	37	16.8	24	2	Q94P09	Q94p09 atlingia ob
15	37	16.8	24	2	Q9SCG5	Q9scg5 mytilaria l
16	37	16.8	24	2	Q9SCG7	Q9scg7 hamamelis j
17	37	16.8	24	2	Q9SCG9	Q9scg9 atlingia ch
18	37	16.8	24	2	Q9SCH1	Q9sch1 liquidambar
19	37	16.8	24	2	Q9SCH3	Q9sch3 liquidambar
20	37	16.8	24	2	Q9SCH5	Q9sch5 semiliquida
21	37	16.8	24	2	Q9SC13	Q9sc13 atlingia ex
22	37	16.8	30	2	Q77211	Q77211 oreochromis
23	37	16.8	35	2	Q9HR54	Q9hrs4 halobacteri
24	37	16.8	37	2	Q9BQ22	Q9bq22 homo sapien
25	37	16.8	41	2	Q8KDX9	Q8kdx9 chlorobium
26	36	16.4	24	2	Q94P10	Q94p10 liquidambar
27	36	16.4	24	2	Q94P11	Q94p11 liquidambar
28	36	16.4	28	2	Q9KQ1	Q9kq1 spingomona
29	36	16.4	28	2	Q9KQ4	Q9kwq4 blastomonas
30	36	16.4	33	2	Q8G3F7	Q8g3f7 galago cras
31	36	16.4	34	2	Q9HF30	Q9hf30 cryptococcu

```

RX PubMed=15024419;
RA Wu M., Sun L.V., Vamathevan J.J., Riegler M., DeBoy R.T.,
RA Brownlie J.C., McGraw E.A., Martin W., Esser C., Ahmadijad N.,
RA Wiegand C., Madupu R., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Kurkin A.S., Kolonay J.F., Nelson W.C., Mohamoud Y., Lee P.,
RA Berry K.J., Young M.B., Utterback T.R., Weidman J.F., Niernan W.C.,
RA Paulsen I.T., Nelson K.E., Tettelin H., O'Neill S.L., Eisen J.A.,
RA "Phylogenomics of the reproductive parasite Wolbachia pipiensis wMel:
RT a streamlined genome overrun by mobile genetic elements.";
RL PLOS Biol. 2:327-341(2004).
DR EMBL; AE017258; AAS14438.1; -.
DR TIGR; WD0747; -.
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 4177 MW; F62611BC059D7523 CRC64;
    Query Match 21.6%; Score 47.5; DB 2; Length 37;
    Best Local Similarity 38.7%; Pred. No. 70;
    Matches 12; Conservative 6; Mismatches 10; Indels 3; Gaps 2;

QY 6 LQTEDSWVPASPDGTGLDPLTVRRHVPVAVWL 36
Db 9 LTSKISWIPVS-SIGMTPYGRHN--DALWLL 36

RESULT 3
Q8NG01 PRELIMINARY; PRT; 32 AA.
ID Q8NG01
AC Q8NG01;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE M3 muscarinic receptor (Fragment).
GN Name=CHRM3;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21855767; PubMed=11867338;
RA Forsythe S.M., Kogut P.C., McConville J.F., Fu Y., McCauley J.A.,
RA Halayko A.J., Liu H.W., Kao A., Fernandes D.J., Bellam S., Fuchs E.,
RA Sinha S., Bell G.I., Camoretti-Mercado B., Solway J.;
RT "Structure and transcription of the human m3 muscarinic receptor
RT gene.";
RL Am. J. Respir. Cell Mol. Biol. 26:298-305(2002).
DR EMBL; AF331838; AAM63959.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON TER
SQ SEQUENCE 32 AA; 3364 MW; 7ECF2F2138B35CC6 CRC64;
    Query Match 17.7%; Score 39; DB 2; Length 32;
    Best Local Similarity 36.8%; Pred. No. 8.1e+02;
    Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 5 FLQTEDSWVPASPDGTGLDP 23
Db 13 FPNISSWIHSPSDGLPP 31

RESULT 4
Q8LGO PRELIMINARY; PRT; 32 AA.
ID Q8LGO
AC Q8LGO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PP1973;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.

```

```

OX NCBI_TaxID=160488;
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfel E.K., Scanlan D., Tran K.,
RA Moazzes A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016781; AAN67588.1; -.
DR TIGR; PP1973; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 32 AA; 3670 MW; 726C36BA63E78739 CRC64;
    Query Match 17.7%; Score 39; DB 2; Length 32;
    Best Local Similarity 46.2%; Pred. No. 8.1e+02;
    Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 12 WVPASPDGTGLDPL 24
Db 13 WAPATVDTAFRPM 25

RESULT 5
Q8VKN1 PRELIMINARY; PRT; 32 AA.
ID Q8VKN1
AC Q8VKN1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein MT0325.
GN OrderedLocusNames=MT0325;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Winn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 194:5479-5490(2002).
DR EMBL; AE006939; AAK44547.1; -.
DR TIGR; MT0325; -.
KW Hypothetical protein.
SQ SEQUENCE 32 AA; 3683 MW; 67353F045A9746B1 CRC64;
    Query Match 17.7%; Score 39; DB 2; Length 32;
    Best Local Similarity 32.3%; Pred. No. 8.1e+02;
    Matches 10; Conservative 4; Mismatches 11; Indels 6; Gaps 2;

QY 14 PASPDGTG---LDPLTVRRHVPVAVWLSDRP 41
Db 4 PHRPHTNEVTEPYTAH---PPLWLTATDP 31

RESULT 6
Q6LDS8 PRELIMINARY; PRT; 39 AA.
ID Q6LDS8

```



```

AC Q61DS8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ca2+ ATPase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=fast twitch muscle;
RX MEDLINE=88169602; PubMed=2965149;
RA Korczak B., Zarain-Herzberg A., Brandl C.J., Ingles C.J., Green N.M.,
RA MacLennan D.H.;
RT "Structure of the rabbit fast-twitch skeletal muscle Ca2+-ATPase
RT gene.";
RL J. Biol. Chem. 263:4813-4819 (1988).
DR EMBL; M20531; AAA31172.1; -.
FT NON_TER 39
SQ SEQUENCE 39 AA; 4366 MW; F2D5C948C7A3A306 CRC64;

Query Match 17.7%; Score 39; DB 2; Length 39;
Best Local Similarity 40.0%; Pred. No. 1e+03;
Matches 10; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 8 TED--SWVPASPDTGLDPLTVRRHV 30
||: ||: ||| ||| ||| |||
Db 9 TEELAYFGVSETTGLTPDQVKRHL 33

RESULT 7
AAA31172
ID AAA31172 PRELIMINARY; PRT; 39 AA.
AC AAA31172;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Ca2+ ATPase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=fast twitch muscle;
RX MEDLINE=88169602; PubMed=2965149;
RA Korczak B., Zarain-Herzberg A., Brandl C.J., Ingles C.J., Green N.M.,
RA MacLennan D.H.;
RT "Structure of the rabbit fast-twitch skeletal muscle Ca2+-ATPase
RT gene.";
RL J. Biol. Chem. 263:4813-4819 (1988).
DR EMBL; M20531; AAA31172.1; -.
FT NON_TER 39
SQ SEQUENCE 39 AA; 4366 MW; F2D5C948C7A3A306 CRC64;

Query Match 17.7%; Score 39; DB 2; Length 39;
Best Local Similarity 40.0%; Pred. No. 1e+03;
Matches 10; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 8 TED--SWVPASPDTGLDPLTVRRHV 30
||: ||: ||| ||| ||| |||
Db 9 TEELAYFGVSETTGLTPDQVKRHL 33

RESULT 8
Q95CH7
ID Q95CH7 PRELIMINARY; PRT; 24 AA.
AC Q95CH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE PsaA (Fragment).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
DE PsaA (Fragment).

```

```

GN Name=psaA;
OS Semiliquidambar cathayensis.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Saxifragales; Hamamelidaceae; Semiliquidambar.
OX NCBI_TaxID=99446;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi S., Huang Y., Zhang Q., Boufford D.E., Parks C.R., Wen J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF304536; AAL13187.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2812 MW; 84E1CC18686377E3 CRC64;

Query Match 17.3%; Score 38; DB 2; Length 24;
Best Local Similarity 38.9%; Pred. No. 7.9e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 24 LTVRRHVPVAVVLLSRDP 41
||: ||: ||| ||| ||| |||
Db 1 MIIRXPEVKILVDRDP 18

RESULT 9
Q95CH9
ID Q95CH9 PRELIMINARY; PRT; 24 AA.
AC Q95CH9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE PsaA (Fragment).
OS Name=psaA;
OG Altingia gracilipes var. serrulata.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Saxifragales; Altingiaceae; Altingia.
OX NCBI_TaxID=166113;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi S., Huang Y., Zhang Q., Boufford D.E., Parks C.R., Wen J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF304535; AAL13185.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2812 MW; 84E1CC18686377E3 CRC64;

Query Match 17.3%; Score 38; DB 2; Length 24;
Best Local Similarity 38.9%; Pred. No. 7.9e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 24 LTVRRHVPVAVVLLSRDP 41
||: ||: ||| ||| ||| |||
Db 1 MIIRXPEVKILVDRDP 18

RESULT 10
Q95CI1
ID Q95CI1 PRELIMINARY; PRT; 24 AA.
AC Q95CI1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE PsaA (Fragment).
OS Name=psaA;
OG Altingia takhtajani.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

```

OC Saxifragales; Altingiaceae; Altingia.

OX NCB1\_TaxID=159571;

RN [1]

RP SEQUENCE FROM N.A.

RA Shi S., Huang Y., Zhang Q., Boufford D.E., Parks C.R., Wen J.;

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF304534; AAL13183.1; -.

GO; GO:0009507; C:chloroplast; IEA.

DR KW Chloroplast.

FT NON TER

SQ SEQUENCE 24 AA; 2812 MW; 84E1CC186B6377E3 CRC64;

Query Match 17.3%; Score 38; DB 2; Length 24;

Best Local Similarity 38.9%; Pred. No. 7.9e+02;

Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 24 LTVRRHVPVAVVLLSRDP 41

Db 1 MIIRSPPEVKILVDRDP 18

RESULT 11

O56554

ID O56554 PRELIMINARY; PRT; 26 AA.

AC O56554;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCB1\_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RA Pan W.S., Wang H.T., Guo H.Z., Wang T.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF040884; AAB95609.1; -.

KW Polyprotein.

FT NON TER

SQ SEQUENCE 26 AA; 2936 MW; E9F8445E75515689 CRC64;

Query Match 17.0%; Score 37.5; DB 2; Length 26;

Best Local Similarity 50.0%; Pred. No. 1e+03;

Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 20 GLDPLTVRRHVPVAVVLL 36

Db 7 GDSFQHLRVHVPVLYIWL 24

RESULT 12

O85UR5

ID O85UR5 PRELIMINARY; PRT; 21 AA.

AC O85UR5;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Photosystem I P700 chlorophyll A apoprotein A1 (fragment).

GN NamepsaA;

OS Strephonema pseudocola.

OC Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC Myrtales; Combretaceae; Strephonema.

OX NCB1\_TaxID=203960;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22466955; PubMed=12579451;

RA Tan F., Shi S., Zhong Y., Gong X., Wang Y.;

RT "Phylogenetic relationships of Combretaceae (Combretaceae) inferred

from plastid, nuclear gene and spacer sequences."

RL J. Plant Res. 115:475-481 (2002).

DR EMBL; AF508246; AAP29780.1; -.

DR GO; GO:0009507; C:chloroplast; IEA.

KW Chloroplast.

FT NON TER

SQ SEQUENCE 21 AA; 2436 MW; B3EB6377E3A17FE6 CRC64;

Query Match 16.8%; Score 37; DB 2; Length 21;

Best Local Similarity 38.9%; Pred. No. 9.3e+02;

Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 24 LTVRRHVPVAVVLLSRDP 41

Db 1 MIIRSPPEVKILVDRDP 18

RESULT 13

O85UR6

ID O85UR6 PRELIMINARY; PRT; 21 AA.

AC O85UR6;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Photosystem I P700 chlorophyll A apoprotein A1 (fragment).

GN Name-psaA;

OS Strephonema manni.

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC Myrtales; Combretaceae; Strephonema.

OX NCB1\_TaxID=203959;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22466955; PubMed=12579451;

RA Tan F., Shi S., Zhong Y., Gong X., Wang Y.;

RT "Phylogenetic relationships of Combretaceae (Combretaceae) inferred

from plastid, nuclear gene and spacer sequences."

RL J. Plant Res. 115:475-481 (2002).

DR EMBL; AF508245; AAP29779.1; -.

DR GO; GO:0009507; C:chloroplast; IEA.

FT NON TER

SQ SEQUENCE 21 AA; 2436 MW; B3EB6377E3A17FE6 CRC64;

Query Match 16.8%; Score 37; DB 2; Length 21;

Best Local Similarity 38.9%; Pred. No. 9.3e+02;

Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 24 LTVRRHVPVAVVLLSRDP 41

Db 1 MIIRSPPEVKILVDRDP 18

RESULT 14

O94P09

ID O94P09 PRELIMINARY; PRT; 24 AA.

AC O94P09;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Psaa (fragment).

GN Name=psaa;

OS Altingia obovata.

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Saxifragales; Altingiaceae; Altingia.

OX NCB1\_TaxID=100935;

RN [1]

RP SEQUENCE FROM N.A.

RA Shi S., Huang Y., Zhang Q., Boufford D.E., Parks C.R., Wen J.;

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF304540; AAL13195.1; -.

Search completed: January 26, 2005, 15:53:20  
Job time : 204 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:35:31 ; Search time 157 Seconds  
(without alignments)

93.681 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_85\_125

Perfect score: 220

Sequence: 1 TKLRFQIQTEDSWVPASPTDG.....DPLTVRRHVPVWVLLSRDP 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 856990

Minimum DB seq length: 0

Maximum DB seq length: 41

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	22.3	37	3	AAB52138 Human sec
2	49	22.3	37	6	ABP99813 Human sec
3	49	22.3	37	6	ABR01305 Human gen
4	49	22.3	37	6	ADA98381 Human sec
5	48.5	22.0	40	8	ABOS5642 Human gen
6	45	20.5	34	3	AAG57081 Arabidops
7	44	20.0	41	4	AAM96628 Human rep
8	44	20.0	41	4	AAU22904 Human pro
9	42.5	19.3	35	2	AAG66872 Human Erb
10	42.5	19.3	32	2	AAW59346 Human Erb
11	42.5	19.3	32	2	AAI14575 Human Erb
12	42.5	19.3	32	4	AAB62075 Human rep
13	42.5	19.3	32	4	AAB48764 Human Erb
14	42	19.1	27	7	ADB84514 MSRV-1B p
15	41.5	18.9	15	3	AAV98932 HLA class
16	41.5	18.9	15	4	AAG89066 HER2/neu
17	41.5	18.9	15	4	AAG88752 HER2/NEU
18	40.5	18.4	32	4	AAM99880 Human exc
19	40.5	18.4	32	4	AAM43679 Human bla
20	40.5	18.4	32	8	ADF71638 Human bla
21	40	18.2	38	4	AAM21828 Peptide #
22	40	18.2	38	4	ABB44199 Peptide #
23	40	18.2	38	4	AAM38152 Peptide #
24	40	18.2	38	4	ABB27077 Protein #
25	40	18.2	38	4	AAM77933 Human bon

#### ALIGNMENTS

##### RESULT 1

AAB52138

ID AAB52138 standard; peptide; 37 AA.

XX AC AAB52138;

XX DT 22-FEB-2001 (first entry)

XX DE Human secreted protein encoded by cDNA #35.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor; vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder; neurological disease; infection; human; secreted protein.

XX OS Homo sapiens.

XX PN WO200061624-A1.

XX PD 19-OCT-2000.

XX PF 06-APR-2000; 2000WO-US008980.

XX PR 09-APR-1999; 99US-0128700P.

XX PR 20-JAN-2000; 2000US-0176930P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

XX DR WPI: 2000-656324/63.

XX DR N-PSDB; AAC96935.

XX CC New nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers.

XX PS Claim 11; Page 439; 478pp; English.

XX CC The invention relates to the isolation of genes AAC96900-C96947 encoding the human secreted proteins AAB52104-B52150. This sequence represents a fragment of the protein encoded by the gene given in the descriptor line. CC The sequence is used as a query sequence for doing BLASTX searches to CC determine homologous sequence to the protein. The genes and proteins are CC useful for preventing, ameliorating or treating medical conditions, e.g. CC by protein or gene therapy. The genes are isolated from a range of human CC tissues disclosed in the specification. The nucleic acids, proteins,

AAM65230 Human bra  
ABG59572 Human liv  
ABG46955 Human pep  
AAM94835 Human rep  
AAB44942 Human sec  
ADK01949 Hepatitis  
ABBS3187 Human ORF  
AAR13378 Adult T-c  
AAR69694 Hepatitis  
AAR89566 Hepatitis  
AAY98841 HLA class  
AAG99011 Her2/neu  
AAG88332 HER2/NEU  
AAB68622 HER-2 B C  
ADH39946 Inhibitor  
AAM14155 Peptide #  
AAM92083 Human dig  
ABB33102 Peptide #  
AAM26563 Peptide #  
ABB27928 Human pep

26 40 18.2 38 4 AAM65230  
27 40 18.2 38 4 ABG59572  
28 40 18.2 38 5 ABG46955  
29 40 18.2 40 4 AAM94835  
30 39 17.7 37 3 AAB44942  
31 39 17.7 40 8 ADK01949  
32 39 17.7 41 5 ABBS3187  
33 38.5 17.5 28 2 AAR13378  
34 38.5 17.5 33 2 AAR69694  
35 38.5 17.5 33 2 AAR89566  
36 38 17.3 15 3 AAY98841  
37 38 17.3 15 4 AAG99011  
38 38 17.3 15 4 AAG88332  
39 38 17.3 19 4 AAB68622  
40 38 17.3 32 8 ADH39946  
41 38 17.3 36 4 AAM14155  
42 38 17.3 36 4 AAM92083  
43 38 17.3 36 4 ABB33102  
44 38 17.3 36 4 AAM26563  
45 38 17.3 36 4 ABB27928

CC antibodies and (ant)agonists are useful in the diagnosis, treatment and  
 CC prevention of: (a) cancer, e.g. breast and ovarian cancer, and other  
 CC cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal  
 CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's  
 CC disease, allergies, autoimmune haemolytic anaemia, autoimmune  
 CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,  
 CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders  
 CC such as myocardial ischaemias; (d) wound healing; (e) neurological  
 CC diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases  
 CC such as viral, bacterial, fungal and parasitic infections  
 XX  
 SQ Sequence 37 AA;

Query Match 22.3%; Score 49; DB 3; Length 37;  
 Best Local Similarity 72.7%; Pred. No. 22;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 26 VRRHVPVAVWL 36  
 :||| ||:|  
 Db 3 LRRHFPALWVL 13

RESULT 2  
 ABP99813  
 ID ABP99813 standard; protein; 37 AA.  
 XX  
 AC ABP99813;  
 XX  
 XX 26-MAR-2003 (first entry)  
 DT  
 XX  
 DE Human secreted protein SEQ ID NO 757.

XX Human; secreted protein; neurotropic; neuroprotective; cytostatic;  
 KW viricide; dermatological; immunosuppressive; antineoplastic; anti-HIV;  
 KW vulnary; antibacterial; antiparkinsonian; anticlinal; antianemic;  
 KW antianthrax; cancer; antineoplastic; hepatotropic; cerebroprotective;  
 KW antineoplastic; antiallergic; antidiabetic; antitumor; anticonvulsant;  
 KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;  
 KW cardiovascular disorder; neurological disease; nephrotropic;  
 KW gene therapy.

XX Homo sapiens.  
 OS  
 XX WO200277186-A2.  
 FN  
 XX 03-OCT-2002.  
 PD  
 XX 26-MAR-2002; 2002WO-US009188.  
 XX  
 XX 27-MAR-2001; 2001US-0278650P.  
 PR  
 XX 12-SEP-2001; 2001US-00950082.  
 PR  
 XX 12-SEP-2001; 2001US-00950083.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Rosen CA, Ruben SM;  
 PI  
 XX WPI; 2003-040583/03.  
 DR  
 XX N-PSDB; ABZ67234.

XX New human secreted proteins encoded by genes contained in cDNA clones  
 PT (e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS,  
 PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or  
 PT West Nile fever.  
 XX  
 XX Claim 1; Page 1490; 2423pp; English.

PS  
 XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the  
 CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,  
 CC treating or ameliorating medical conditions e.g. by protein or gene  
 CC therapy. The genes are isolated from a range of human tissues disclosed  
 CC in the specification. The nucleic acids, proteins, antibodies and  
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:

CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the  
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 CC lung or urogenital; (b) immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as  
 CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.  
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,  
 CC bacterial, fungal and parasitic infections  
 XX  
 SQ Sequence 37 AA;

Query Match 22.3%; Score 49; DB 6; Length 37;  
 Best Local Similarity 72.7%; Pred. No. 22;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 26 VRRHVPVAVWL 36  
 :||| ||:|  
 Db 3 LRRHFPALWVL 13

RESULT 3  
 ABR01305  
 ID ABR01305 standard; peptide; 37 AA.  
 XX  
 AC ABR01305;  
 XX  
 XX 12-MAY-2003 (first entry)  
 DT  
 XX  
 DE Human gene 359-encoded secreted protein HTXMZ07, SEQ ID NO:786.

XX Human; secreted protein; cancer; tumour; hyperproliferative disorder;  
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;  
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;  
 KW drug screening; chromosome identification; chromosome mapping;  
 KW cytostatic; gene therapy; antineoplastic; immunomodulator; anti-HIV;  
 KW antianemic; vulnary; chromosome 3p21.31.

XX Homo sapiens.  
 OS  
 XX WO200277013-A2.  
 FN  
 XX 03-OCT-2002.  
 PD  
 XX 26-MAR-2002; 2002WO-US009370.  
 XX  
 XX 27-MAR-2001; 2001US-0278650P.  
 PR  
 XX 12-SEP-2001; 2001US-00950082.  
 PR  
 XX 12-SEP-2001; 2001US-00950083.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Rosen CA, Ruben SM;  
 PI  
 XX WPI; 2003-040578/03.  
 DR  
 XX N-PSDB; ABZ73639.

XX New human secreted proteins and nucleic acids, useful for detecting or  
 PT treating cancer or other hyperproliferative disorders, autoimmune  
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.  
 XX  
 XX Claim 13; Page 1491; 2474pp; English.

XX ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted  
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.  
 CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins are thought to be involved in biological activities  
 CC associated with cellular signalling, cellular differentiation, cell  
 CC migration, prohormone activation and neurotransmitter activity. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody

CC fragments specific for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing or treating cancers or other  
 CC hyperproliferative disorders. Additionally, the secreted proteins and  
 CC their nucleic acids may also be used in the treatment of autoimmune  
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS  
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote  
 CC wound healing. Nucleic acids of the invention may be used for chromosome  
 CC identification, chromosome mapping, in gene therapy, for identifying  
 CC individuals from minute biological samples, as hybridisation probes, and  
 CC as molecular weight markers. The present sequence represents a human  
 CC secreted protein of the invention

XX SQ Sequence 37 AA;

Query Match 22.3%; Score 49; DB 6; Length 37;  
 Best Local Similarity 72.7%; Pred. No. 22;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 26 VRHVPVAVWL 36  
 :||| ||:|  
 Db 3 LRRHFPALMWL 13

RESULT 4  
 ADA98381  
 ID ADA98381 standard; protein; 37 AA.

XX AC ADA98381;

DT 20-NOV-2003 (first entry)

XX DE Human secreted protein sequence #222.

XX KW human; secreted protein; cardiovascular disorder; arrhythmia;  
 KW atherosclerosis; stroke; endocarditis; congestive heart failure;  
 KW rheumatic heart disease; cardiomyopathy; haemorrhoids; varicose veins;  
 KW migraine; thrombosis; neural disorder; immune system disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; renal disorder; proliferative disorder; cancer.

XX OS Homo sapiens.

XX FN WO2003004623-A2.

XX PD 16-JAN-2003.

XX PF 26-MAR-2002; 2002WO-US009922.

XX PR 27-MAR-2001; 2001US-0278650P.

XX PR 12-SEP-2001; 2001US-00950082.

XX PR 12-SEP-2001; 2001US-00950083.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX WPI; 2003-247946/24.

XX PT New human secreted polypeptide and nucleic acid molecules, useful for  
 PT diagnosing, preventing, prognosticating or treating cardiovascular  
 PT disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or  
 PT thrombosis).

XX PS Claim 1; SEQ ID NO 490; 1572pp; English.

XX CC The invention comprises the amino acid and coding sequence of human  
 CC secreted proteins. The DNA and protein sequences of the invention are  
 CC useful in the treatment of cardiovascular disorders, such as: arrhythmia,  
 CC atherosclerosis, stroke, endocarditis, congestive heart failure,  
 CC rheumatic heart disease, cardiomyopathy, haemorrhoids, varicose veins,  
 CC migraine, or thrombosis. The DNA and protein sequences may also be used  
 CC for treating or preventing: neural disorders, immune system disorders,  
 CC muscular disorders, reproductive disorders, gastrointestinal disorders,

CC pulmonary disorders, renal disorders, proliferative disorders and/or  
 CC cancerous diseases. The present amino acid sequence represents a human  
 CC secreted protein of the invention. NOTE: The present sequence is shown on  
 CC the WIPO website.

XX SQ Sequence 37 AA;

Query Match 22.3%; Score 49; DB 6; Length 37;  
 Best Local Similarity 72.7%; Pred. No. 22;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 26 VRHVPVAVWL 36  
 :||| ||:|  
 Db 3 LRRHFPALMWL 13

RESULT 5

ABO55642  
 ID ABO55642 standard; protein; 40 AA.

XX AC ABO55642;

DT 29-JUL-2004 (first entry)

XX DE Human genome derived single exon protein #1876.

XX KW Human; gene expression; single exon probe; microarray;  
 KW alternative splicing event; genomic alteration.

XX OS Homo sapiens.

XX FN US2003194704-A1.

XX PD 16-OCT-2003.

XX PF 03-APR-2002; 2002US-00029386.

XX PR 03-APR-2002; 2002US-00029386.

XX PA (PENN/) PENN S G.

XX PA (RANK/) RANK D R.

XX PA (HANZ/) HANZEL D K.

XX PI Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

XX PT New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.

XX PS Claim 45; SEQ ID NO 29276; 80pp; English.

XX CC The invention relates to a nucleic acid probe for measuring human gene  
 CC expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridises under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing

human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docid=20030194704

XX SQ Sequence 40 AA;

Query Match 22.0%; Score 48.5; DB 8; Length 40;

Best Local Similarity 55.0%; Pred. No. 29;

Matches 1; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Qy 12 WVPASPDGDLPLVRHVP 31

| : | | | | | | | : |

Db 18 WLFASAD-NLDPLTLRNTNP 36

RESULT 6

AAG57081  
ID AAG57081 standard; protein; 34 AA.

XX AC AAG57081;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 73511.

XX KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 20-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 22-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.



PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147182P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 13-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 07-OCT-1999; 99US-0157865P.  
PR 08-OCT-1999; 99US-0158029P.  
PR 12-OCT-1999; 99US-0158232P.  
PR 13-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 20.5%; Score 45; DB 3; Length 34;  
Best Local Similarity 55.6%; Pred. NO. 74;  
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 7 QTEDSWPASPDTGLDPL 24  
|: ||| | || ||  
Db 19 QSRDSWFP--PDRLRLPL 34

## RESULT 7

AAM96628

ID AAM96628 standard; protein; 41 AA.

XX AC AAM96628;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen SEQ ID NO: 5286.

KW Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001339.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 11-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

```
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232357P.
PR 14-SEP-2000; 2000US-0232388P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX N-PSDB; AAL02598.
XX WPI: 2001-465570/50.
XX CC number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a protein of the invention
XX
XX Claim 11; SEQ ID NO 5286; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a protein of the invention
XX
XX Sequence 41 AA;
XX
XX Query Match 20.0%; Score 44; DB 4; Length 41;
XX Best Local Similarity 35.5%; Pred. No. 1.3e+02;
XX Matches 11; Conservative 7; Mismatches 9; Indels 4; Gaps 2;
XX
XX QY 2 KLRFLQTEDSWVPASPTGLDPLTVR---RH 29
XX ::::: |:::| |:::|
XX Db 2 QIRFVAIEANWIPAQL-CGLDSHHKSTRH 31
XX
XX RESULT 8
XX AAU22904
XX ID AAU22904 standard; protein; 41 AA.
XX AC AAU22904;
XX XX
XX DT 17-DEC-2001 (first entry)
XX
XX DE Human prostate cancer antigen, Seq ID No 423.
XX
XX Human; prostate cancer antigen; cytostatic; uropathic; diagnostic;
XX reproductive system; chromosomal marker; forensic; urinary disorder;
XX KW
```

KW chronic nephritis; blood-related disorder; thrombosis.  
XX Homo sapiens.  
OS  
PN WO200155316-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001328.  
PP  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218230P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.



CC inflammatory, angiogenic and immunologic disorders

XX Sequence 32 AA;

Query Match 19.3%; Score 42.5; DB 2; Length 32;  
Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30  
:||||:|||||  
Db 14 LPASPETHLDML---RHL 28

RESULT 11

AAAY14575  
ID AAY14575 standard; peptide; 32 AA.

AC AAY14575;

XX 14-SEP-1999. (first entry)

XX Human erbB2 receptor-derived epitope 7C2/7F3.

XX Domain; human; erbB2; receptor; antibody; overexpression; anthracycline;  
XX chemotherapeutic agent; tumour; cancer; breast; squamous cell; cervical;  
KW non-small-cell lung; gastrointestinal; pancreatic; glioblastoma; ovarian;  
KW liver; bladder; hepatoma; colon; colorectal; endometrial carcinoma;  
KW salivary gland; kidney; prostate; vulva; thyroid; hepatic carcinoma;  
KW head-and-neck; epitope.

XX Homo sapiens.

OS

XX WO9931140-A1.

XX 24-JUN-1999.

XX 10-DEC-1998; 98WO-US026266.

XX 12-DEC-1997; 97US-0069346P.

XX (GETH ) GENENTECH INC.

XX Shak S, Paton VE;

XX WPI; 1999-395164/33.

XX New combination of an anti-ErbB2 antibody and a chemotherapeutic agent.

XX Disclosure; Page 35; 42pp; English.

XX This sequence represents the 7C2/7F3 epitope from domain 1 of the human  
XX erbB2 receptor protein. ErbB2 protein is used for raising antibodies for  
XX treating a human with a disorder characterized by overexpression of the  
XX ErbB2 receptor, using a combination of an anti-ErbB2 antibody and a  
XX chemotherapeutic agent other than an anthracycline derivative. Disorders  
XX characterised by overexpression of the ErbB2 receptor include benign or  
XX malignant tumours. The cancer may be breast, squamous cell, (non)-small-  
XX cell lung, gastrointestinal, pancreatic, glioblastoma, cervical, ovarian,  
XX liver, bladder, hepatoma, colon, colorectal, endometrial carcinoma,  
XX salivary gland carcinoma, kidney, prostate, vulval or thyroid cancer,  
XX hepatic carcinoma and various forms of head-and-neck cancer

XX Sequence 32 AA;

Query Match 19.3%; Score 42.5; DB 2; Length 32;  
Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30  
:||||:|||||  
Db 14 LPASPETHLDML---RHL 28

RESULT 12

AAB62075  
ID AAB62075 standard; peptide; 32 AA.

XX AAB62075;

XX 29-MAY-2001 (first entry)

XX ErbB2 7C2/7F3 epitope.

XX ErbB2 receptor; anti-ErbB2 antibody; tumour; cancer; breast cancer;  
KW leukemia; cytostatic; carcinoma; epitope.

XX Homo sapiens.

XX WO200115730-A1.

XX 08-MAR-2001.

XX 25-AUG-2000; 2000WO-US023391.

XX 27-AUG-1999; 99US-0151018P.

XX 23-JUN-2000; 2000US-0213822P.

XX (GETH ) GENENTECH INC.

XX Baughman SA, Shak S;

XX WPI; 2001-235058/24.

XX Treating a human patient with a disorder characterized by overexpression  
PT of ErbB2 receptor such as a tumor or cancer (e.g. metastatic breast  
PT carcinoma) comprises administering several dosages of an anti-ErbB2  
PT antibody.

XX Disclosure; Page 64; 71pp; English.

XX The invention relates to treating cancer in a human susceptible to or  
CC diagnosed with a disorder characterized by overexpression of ErbB2  
CC receptor. The method comprises administering an initial dose of at least  
CC approx.5 mg/kg of the anti-ErbB2 antibody, and administering subsequent  
CC doses of the antibody in approximately the same amount or less than the  
CC initial dose. The method is useful for treating a disorder characterized  
CC by overexpression of ErbB2 receptor, particularly a benign or malignant  
CC tumour, or a cancer. The cancer is selected from breast cancer  
CC (particularly metastatic breast carcinoma), leukemia, squamous cell  
CC cancer, small-cell lung cancer, non-small cell lung cancer, pancreatic  
CC cancer, gastrointestinal cancer, glioblastoma, cervical cancer, ovarian  
CC cancer, liver cancer, bladder cancer, hepatoma, colon cancer, colorectal  
CC cancer, endometrial carcinoma, salivary gland carcinoma, kidney cancer,  
CC liver cancer, prostate cancer, vulval cancer, thyroid cancer, hepatic  
CC carcinoma and various types of head and neck cancer. The present sequence  
CC represents the 7C2/7F3 peptide epitope of ErbB2, recognised by MAb8 7C2  
CC and 7F3 as determined by deletion mapping

XX Sequence 32 AA;

Query Match 19.3%; Score 42.5; DB 4; Length 32;  
Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30  
:||||:|||||  
Db 14 LPASPETHLDML---RHL 28

RESULT 13

AAB48764  
ID AAB48764 standard; protein; 32 AA.

XX AAB48764;

XX 09-MAR-2001 (first entry)



Search completed: January 26, 2005, 15:49:52  
Job time : 163 secs

ID AAY98932 standard; peptide; 15 AA.  
XX  
AC AAY98932;  
XX  
DT 07-AUG-2000 (first entry)  
XX  
DE HLA class II binding antigen epitope peptide #121.  
XX  
KW Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;  
KW immune response; chronic viral disease; cancer; autoimmune disease;  
KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;  
KW allograft rejection; allergy; Lyme disease; hepatitis; prostate cancer;  
KW glomerulonephritis; food hypersensitivity; malaria.  
XX  
OS Unidentified.  
XX  
PN WO9961916-A1.  
XX  
PD 02-DEC-1999.  
XX  
PF 28-MAY-1999; 99WO-US012066.  
XX  
PR 29-MAY-1998; 98US-0087192P.  
XX  
PA (EPIM-) EPIMMUNE INC.  
XX  
PI Sette A, Southwood S, Sidney J;  
XX  
PI WPI; 2000-097143/08.  
XX  
DR  
XX  
PT New compositions containing immunogenic peptide epitopes for various HLA  
PT class II DR molecules useful for inducing helper T cell response.  
XX  
PS Claim 1; Page 42; 60pp; English.  
XX  
CC The present invention relates to a new pharmaceutical composition  
CC comprising a unit dose form of a peptide, or analogue, comprising an  
CC epitope selected from those represented by peptides AAY98812-Y99339 which  
CC are derived from various antigens for various human leucocyte antigen  
CC class DR molecules, representative of the world wide population. The  
CC peptide/analogue binds to an HLA class II molecule at an IC-50 of less  
CC than or equal to 1,000 nM. The pharmaceutical can be used to induce a  
CC helper T cell response. The pharmaceutical focuses the immune response  
CC towards selected determinants and could therefore be used in cases of  
CC chronic viral diseases and cancer. Examples of diseases that can be  
CC treated using the peptide containing pharmaceutical include autoimmune  
CC diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia  
CC gravis), allograft rejection, allergies, Lyme disease, hepatitis, post-  
CC streptococcal endocarditis or glomerulonephritis and food  
CC hypersensitivities. The peptide epitopes can be used to enhance immune  
CC responses against other immunogens administered with the peptides.  
CC Diseases which can be treated using immunogenic mixtures include prostate  
CC cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical  
CC carcinoma, lymphoma, and condyloma acuminatum. The peptides may also be  
CC used to make monoclonal antibodies useful as potential diagnostic or  
CC therapeutic agents. The peptides may also be useful as diagnostic  
CC reagents, for example, to determine the susceptibility of an individual  
CC to a treatment regimen. Also, the peptides may be used to predict which  
CC individuals will be at substantial risk of developing chronic infection.  
CC The selection of appropriate T and B cell epitopes should allow the  
CC development of epitope based vaccines particularly towards conserved  
CC epitopes of pathogens which are characterized by high sequence  
CC variability such as HIV, HCV and Malaria  
XX  
SQ Sequence 15 AA;  
  
Query Match 18.9%; Score 41.5; DB 3; Length 15;  
Best Local Similarity 58.8%; Pred. No. 86;  
Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;  
  
Qy 13 VPASPTGLDPLVRRH 29  
:|||||:  
Db 2 LPASPTHLDML---RH 15

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:53:27 ; Search time 144 Seconds  
(without alignments)  
102.867 Million cell updates/sec

Title: US-10-057-813-14\_COPY\_85\_125

Perfect score: 220  
Sequence: 1 TKLRFQTEDSWPASPDTG.....DPLTVRRHVPVWVLLSRDP 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 391627

Minimum DB seq length: 0  
Maximum DB seq length: 41

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48.5	22.0	40	14	US-10-029-386-29276
2	44	20.0	41	10	US-09-764-891-5286
3	44	20.0	41	14	US-10-091-572-423
4	43	19.5	41	15	US-10-424-599-213057
5	42.5	19.3	32	14	US-10-356-824-2
6	42.5	19.3	32	14	US-10-406-925-2
7	42.5	19.3	32	14	US-10-429-519-2
8	42.5	19.3	32	15	US-10-356-824-2
9	42.5	19.3	32	15	US-10-600-152-2
10	42.5	19.3	32	17	US-10-909-998-2
11	42	19.1	37	17	US-10-425-115-334035
12	42	19.1	40	17	US-10-425-115-249550
13	41.5	18.9	15	15	US-10-149-138-3951

14	41.5	18.9	15	15	US-10-149-138-4508	Sequence 4508, Ap
15	41.5	18.9	15	16	US-10-149-138-3951	Sequence 3951, Ap
16	41.5	18.9	15	16	US-10-149-138-4508	Sequence 4508, Ap
17	41	18.6	25	9	US-09-729-835-96	Sequence 96, Appl
18	41	18.6	25	15	US-10-373-809-96	Sequence 96, Appl
19	40.5	18.4	32	14	US-10-080-254-73	Sequence 73, Appl
20	40.5	18.4	32	14	US-10-242-355-617	Sequence 617, App
21	40.5	18.4	40	15	US-10-424-599-187800	Sequence 187800,
22	40	18.2	37	17	US-10-425-115-228363	Sequence 228363,
23	40	18.2	38	9	US-09-864-761-42375	Sequence 42375, A
24	40	18.2	38	9	US-09-843-676-27	Sequence 27, Appl
25	40	18.2	38	9	US-09-766-253-27	Sequence 27, Appl
26	40	18.2	38	10	US-09-438-486-27	Sequence 27, Appl
27	40	18.2	38	14	US-10-053-758-27	Sequence 27, Appl
28	40	18.2	38	14	US-10-054-295-27	Sequence 27, Appl
29	40	18.2	38	14	US-10-054-611-27	Sequence 27, Appl
30	40	18.2	38	14	US-10-325-810-217	Sequence 217, App
31	40	18.2	38	17	US-10-877-124-217	Sequence 217, App
32	40	18.2	38	17	US-10-877-022-217	Sequence 217, App
33	40	18.2	40	10	US-09-764-891-3493	Sequence 3493, Ap
34	39	17.7	11	15	US-10-149-138-570	Sequence 570, App
35	39	17.7	11	15	US-10-149-138-1921	Sequence 1921, Ap
36	39	17.7	11	16	US-10-149-138-570	Sequence 570, App
37	39	17.7	11	16	US-10-149-138-1921	Sequence 1921, Ap
38	39	17.7	33	12	US-09-855-604-715	Sequence 715, App
39	39	17.7	40	10	US-10-425-115-308261	Sequence 308261,
40	39	17.7	41	10	US-09-826-734-106	Sequence 106, App
41	38	17.3	11	15	US-10-149-138-1515	Sequence 1515, Ap
42	38	17.3	11	15	US-10-149-138-2164	Sequence 2164, Ap
43	38	17.3	11	16	US-10-149-138-1515	Sequence 1515, Ap
44	38	17.3	11	16	US-10-149-138-2164	Sequence 2164, Ap
45	38	17.3	15	15	US-10-149-138-3741	Sequence 3741, Ap

#### ALIGNMENTS

#### RESULT 1

```
US-10-029-386-29276
; Sequence 29276, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29276
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHRS.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
```

Query Match 22.0%; Score 48.5; DB 14; Length 40;

Best Local Similarity 55.0%; Pred. No. 55;

Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Qy 12 WVPASPDGDLPLTVRRHVP 31

Db 18 WLFASAD-NLDPPLTLRPNT 36

RESULT 2  
US-09-764-891-5286  
; Sequence 5286, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5286  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-891-5286

Query Match 20.0%; Score 44; DB 10; Length 41;  
Best Local Similarity 35.5%; Pred. No. 2.2e+02;  
Matches 11; Conservative 7; Mismatches 9; Indels 4; Gaps 2;  
QY 2 KLRFLQTESWVSPDPTGLDPLTVR---RH 29  
Db 2 QIRFVAIEANWIPQAL-CGLDLSHKHSSTRH 31

RESULT 3  
US-10-091-572-423  
; Sequence 423, Application US/10091572  
; Publication No. US20030054373A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PALL8C1  
; CURRENT APPLICATION NUMBER: US/10/091,572  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: 09/764,850  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20

```

RESULT 5
US-10-356-824-2
, Sequence 2, Application US/10356824
, Publication No. US20030147884A1
, GENERAL INFORMATION:
, APPLICANT: Shak, Steve
, APPLICANT: Paton, Virginia
, TITLE OF INVENTION: TREATMENT WITH ANTI-ERBB2 ANTIBODIES
, FILE REFERENCE: P1256R1
, CURRENT APPLICATION NUMBER: US/10/356,824
, CURRENT FILING DATE: 2003-02-03
, PRIOR APPLICATION NUMBER: US/09/208,649
, PRIOR FILING DATE: 1998-12-10
, PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/069,346
, PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-12
, NUMBER OF SEQ ID NOS: 9
, SEQ ID NO 2
, LENGTH: 32

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-356-824-2

Query Match          19.3%; Score 42.5; DB 14; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDTCGLDPLTVRRHV 30
    :||||:| | | | |
Db 14 LPASPETHLDML---RHL 28

RESULT 6
US-10-406-925-2
; Sequence 2, Application US/10406925
; Publication No. US20030170234A1
; GENERAL INFORMATION:
; APPLICANT: Shak, Steve
; APPLICANT: Paton, Virginia
; TITLE OF INVENTION: TREATMENT WITH ANTI-ERBB2 ANTIBODIES
; FILE REFERENCE: P1256R3
; CURRENT APPLICATION NUMBER: US/10/406,925
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/209,023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/069,346
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-406-925-2

Query Match          19.3%; Score 42.5; DB 14; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDTCGLDPLTVRRHV 30
    :||||:| | | | |
Db 14 LPASPETHLDML---RHL 28

RESULT 7
US-10-429-519-2
; Sequence 2, Application US/10429519
; Publication No. US20030170235A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Robert
; TITLE OF INVENTION: TREATMENT WITH ANTI-ERBB2 ANTIBODIES
; FILE REFERENCE: P1757R1
; CURRENT APPLICATION NUMBER: US/10/429,519
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/568,322
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US 60/134,085
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-429-519-2

Query Match          19.3%; Score 42.5; DB 14; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDTCGLDPLTVRRHV 30
    :||||:| | | | |
Db 14 LPASPETHLDML---RHL 28

```

; FILE REFERENCE: P1256R3  
; CURRENT APPLICATION NUMBER: US/10/909,998  
; CURRENT FILING DATE: 2004-08-02  
; PRIOR APPLICATION NUMBER: US/09/209,023  
; PRIOR FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: US 60/069,346  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 2  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-909-998-2

Query Match 19.3%; Score 42.5; DB 17; Length 32;  
Best Local Similarity 55.6%; Pred. No. 2.7e+02;  
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPTGLDPLTVRRHV 30  
:||||:|||||  
Db 14 LPASPETHDML---RHL 28

RESULT 11  
US-10-425-115-334035  
; Sequence 334035, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 334035  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_67752C.1.pep  
US-10-425-115-334035

Query Match 19.1%; Score 42; DB 17; Length 37;  
Best Local Similarity 53.8%; Pred. No. 3.6e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LRFLOTEDSWVPA 15  
:||||:|||||  
Db 25 IRLLETPDAWTA 37

RESULT 12  
US-10-425-115-249550  
; Sequence 249550, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 249550  
; LENGTH: 40  
; TYPE: PRT

; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_159174C.1.pep  
US-10-425-115-249550

Query Match 19.1%; Score 42; DB 17; Length 40;  
Best Local Similarity 36.4%; Pred. No. 4e+02;  
Matches 8; Conservative 5; Mismatches 7; Indels 2; Gaps 1;

QY 9 EDSWVPASPTGLDPLTVRRHV 30  
:||||:|||||  
Db 6 QETWM--DPAIGLTPYDVHKKV 25

RESULT 13  
US-10-149-138-3951  
; Sequence 3951, Application US/10149138  
; Publication No. US20040018971A1  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Cellis, Esteban  
; APPLICANT: Keogh, Elissa  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to  
; TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
; FILE REFERENCE: 2060.0140001  
; CURRENT APPLICATION NUMBER: US/10/149,138  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: PCT/US00/33591  
; PRIOR FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: US 09/458,299  
; PRIOR FILING DATE: 1999-12-11  
; NUMBER OF SEQ ID NOS: 4641  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 3951  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-3951

Query Match 18.9%; Score 41.5; DB 15; Length 15;  
Best Local Similarity 58.8%; Pred. No. 1.5e+02;  
Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPTGLDPLTVRRHV 29  
:||||:|||||  
Db 2 LPASPETHDML---RH 15

RESULT 14  
US-10-149-138-4508  
; Sequence 4508, Application US/10149138  
; Publication No. US20040018971A1  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Cellis, Esteban  
; APPLICANT: Keogh, Elissa  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to  
; TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
; FILE REFERENCE: 2060.0140001  
; CURRENT APPLICATION NUMBER: US/10/149,138  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: PCT/US00/33591  
; PRIOR FILING DATE: 2000-12-11

; PRIOR APPLICATION NUMBER: US 09/458,299  
; PRIOR FILING DATE: 1999-12-11  
; NUMBER OF SEQ ID NOS: 4641  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4508  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial Peptide  
US-10-149-138-4508

Query Match 18.9%; Score 41.5; DB 15; Length 15;  
Best Local Similarity 58.8%; Pred. No. 1.5e+02;  
Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

Qy 13 VPASPDTCGLDPLTVRRH 29  
:||||:| | | | |  
Db 2 LPASPETHLDML---RH 15

## RESULT 15

US-10-149-138-3951  
; Sequence 3951, Application US/10149138  
; Publication No. US20040121946A9  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Eteban  
; APPLICANT: Keogh, Elissa  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to  
; TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
; FILE REFERENCE: 2060.0140001  
; CURRENT APPLICATION NUMBER: US/10/149,138  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: PCT/US00/33591  
; PRIOR FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: US 09/458,299  
; PRIOR FILING DATE: 1999-12-11  
; NUMBER OF SEQ ID NOS: 4641  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3951  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-3951

Query Match 18.9%; Score 41.5; DB 16; Length 15;  
Best Local Similarity 58.8%; Pred. No. 1.5e+02;  
Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

Qy 13 VPASPDTCGLDPLTVRRH 29  
:||||:| | | | |  
Db 2 LPASPETHLDML---RH 15

Search completed: January 26, 2005, 16:05:52  
Job time : 145 secs

Result No.	Score			Query Match	Length	DB	ID	Description
	1	2	3					
1	45.5	20.7	40	4	US-09-270-767-41398	Sequence 41398, A		
2	45.5	20.7	40	3	US-09-270-767-56614	Sequence 56614, A		
3	45	20.5	17	3	US-09-174-465D-6	Sequence 6, Appli		
4	45	20.5	17	3	US-09-599-564A-6	Sequence 2, Appli		
5	42.5	19.3	32	4	US-09-648-067A-2	Sequence 96, Appli		
6	41	18.6	25	4	US-09-257-179-96	Sequence 27, Appl		
7	40	18.2	38	3	US-08-851-843A-27	Sequence 217, App		
8	40	18.2	38	3	US-08-974-549A-217	Sequence 27, Appl		
9	40	18.2	38	3	US-08-854-050-27	Sequence 27, Appl		
10	40	18.2	38	3	US-09-430-323-27	Sequence 27, Appl		
11	40	18.2	38	4	US-09-402-181B-217	Sequence 217, App		
12	40	18.2	38	2	US-09-721-456-217	Sequence 217, App		
13	39	17.7	35	2	US-08-726-306A-39	Sequence 39, Appl		
14	39	17.7	38	1	US-09-270-767-56715	Sequence 56715, A		
15	38.5	17.5	28	1	US-07-596-081A-14	Sequence 14, Appl		
16	38.5	17.5	33	1	US-08-086-428B-145	Sequence 145, App		
17	38.5	17.5	33	2	US-08-468-570-145	Sequence 145, App		
18	38.5	17.5	33	2	US-08-290-668A-249	Sequence 249, App		
19	38.5	17.5	33	4	US-08-466-601A-145	Sequence 145, App		
20	38.5	17.5	33	5	PCR-US95-10398-249	Sequence 249, App		
21	37.5	17.0	34	4	US-09-645-470-3	Sequence 3, Appli		
22	37.5	17.0	34	4	US-09-355-040-20	Sequence 20, Appl		
23	37	16.8	25	3	US-08-612-973-55	Sequence 55, Appl		
24	37	16.8	25	3	US-08-927-597-55	Sequence 55, Appl		
25	37	16.8	25	3	US-08-851-843A-26	Sequence 26, Appl		
26	37	16.8	39	3	US-08-974-549A-216	Sequence 216, App		
27	37	16.8	39	3	US-08-854-050-26	Sequence 26, Appl		





RESULT 7  
US-08-851-843A-27  
; Sequence 27, Application US/08851843A  
; Patent No. 6093809  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: NO. 6093809a1 Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/851,843A  
; FILING DATE: 06-MAY-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-851-843A-27

Query Match 18.2%; Score 40; DB 3; Length 38;  
Best Local Similarity 52.9%; Pred. No. 1.1e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;  
Qy 4 RFLQT----EDSWTPAS 16  
|||:|  
Db 18 RFLRTAEXKNDGWVPIS 34

RESULT 8  
US-08-974-549A-217  
; Sequence 217, Application US/08974549A  
; Patent No. 6166178

; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/974,549A  
; FILING DATE: 19-NOV-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/911,312  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US97/17618  
; FILING DATE: 01-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US97/17885  
; FILING DATE: 01-OCT-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph Ted  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002610US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 217:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-974-549A-217  
Query Match 18.2%; Score 40; DB 3; Length 38;

Best Local Similarity 52.9%; Pred. No. 1.1e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 4 RFLQT-----EDSWVPAS 16  
|||:| | | | |  
Db 18 RFLRTTAEKNDGWVPIS 34

RESULT 9  
US-08-854-050-27  
; Sequence 27, Application US/08854050  
; Patent No. 6261836  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: No. 6261836el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/854,050  
; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; US-08-854-050-27

Query Match 18.2%; Score 40; DB 3; Length 38;  
Best Local Similarity 52.9%; Pred. No. 1.1e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 4 RFLQT-----EDSWVPAS 16  
|||:| | | | |  
Db 18 RFLRTTAEKNDGWVPIS 34

QY 4 RFLQT-----EDSWVPAS 16  
|||:| | | | |  
Db 18 RFLRTTAEKNDGWVPIS 34

RESULT 10  
US-09-430-323-27  
; Sequence 27, Application US/09430323  
; Patent No. 6309867  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: No. 6309867el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/430,323  
; FILING DATE: 29-Oct-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-09-430-323-27

Query Match 18.2%; Score 40; DB 3; Length 38;  
Best Local Similarity 52.9%; Pred. No. 1.1e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 4 RFLQT-----EDSWVPAS 16  
|||:| | | | |  
Db 18 RFLRTTAEKNDGWVPIS 34

RESULT 11  
US-09-402-181B-217  
; Sequence 217, Application US/09402181B  
; Patent No. 6610839  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin B.  
; Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 633  
; CORRESPONDENCE ADDRESS:  
; STREET: Townsend and Townsend and Crew LLP  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/402,181B  
; FILING DATE: 29-Sep-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/911,312  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: WO PCT/US97/17885  
; FILING DATE: 01-OCT-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ausenhus, Scott L.  
; REGISTRATION NUMBER: 42,271  
; REFERENCE/DOCKET NUMBER: 015389-002620US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 217:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 217:  
US-09-402-181B-217

Query Match 18.2%; Score 40; DB 4; Length 38;  
Best Local Similarity 52.9%; Pred. No. 1.1e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

Qy 4 RFLQT-----EDSWVPAS 16  
Db 18 RFLRTTAEKNDGWVPIS 34

RESULT 12  
US-09-721-456-217  
; Sequence 217, Application US/09721456  
; Patent No. 6617110  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin B.  
; Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; STREET: Townsend and Townsend and Crew LLP  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/721,456  
; FILING DATE: 22-No. 6617110-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/974,549A  
; FILING DATE: 19-NOV-1997  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/911,312  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: US 08/915,503  
; APPLICATION NUMBER: WO PCT/US97/17618  
; FILING DATE: 01-OCT-1997  
; APPLICATION NUMBER: WO PCT/US97/17885  
; FILING DATE: 01-OCT-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph Ted  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002610US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 217:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 217:  
US-09-721-456-217

Query Match 18.2%; Score 40; DB 4; Length 38;

Best Local Similarity 52.9%; Pred. No. 1.1e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 4 RFLQT-----ESWVPAS 16  
| | | | | | | | | |  
Db 18 RFLRTTAEKNDGWVPIS 34

## RESULT 13

US-08-726-306A-39  
; Sequence 39, Application US/08726306A  
; Patent No. 5958684  
; GENERAL INFORMATION:  
; APPLICANT: van Leeuwen, Frederik Willem  
; APPLICANT: Burbach, Johannes Peter Henri  
; APPLICANT: Grosveld, Franklin G.  
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS  
; NUMBER OF SEQUENCES: 189  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1 Financial Center  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/726,306A  
; FILING DATE: 02-Oct-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 95/20080.4  
; FILING DATE: 02-Oct-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/009,832  
; FILING DATE: 01-Jan-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams, Ph.D. Kathleen M.  
; REGISTRATION NUMBER: 34,380  
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 345-9100  
; TELEFAX: (617) 345-9111  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-726-306A-39

Query Match 17.7%; Score 39; DB 2; Length 35;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TKLRLQTEDSWVP 14  
| | | | | | | | | |  
Db 22 TKLFFTHRHSMNP 35

## RESULT 14

US-09-270-767-56715  
; Sequence 56715, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 56715  
; LENGTH: 38  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-56715

Query Match 17.7%; Score 39; DB 4; Length 38;  
Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 12 WVPASPT 19  
| | | | | | | | | |  
Db 11 WIPETPDT 18

## RESULT 15

US-07-596-081A-14  
; Sequence 14, Application US/07596081A  
; Patent No. 5194586  
; GENERAL INFORMATION:  
; APPLICANT: Maeda, Yoshiaki  
; APPLICANT: Shiraki, Hiroshi  
; APPLICANT: Washitani, Yukiko  
; APPLICANT: Kuroda, Naotaka  
; APPLICANT: Yamada, Kyoko  
; APPLICANT: Oka, Kiichiro  
; APPLICANT: Namba, Toshihiro  
; TITLE OF INVENTION: Peptides and Use Thereof  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eric S. Spector  
; STREET: P.O. Box 2266 Eads Station  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/596,081A  
; FILING DATE: 19901011  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 266983/1989  
; FILING DATE: 13-OCT-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spector, Eric S.  
; REGISTRATION NUMBER: 22495  
; REFERENCE/DOCKET NUMBER: 513904  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-415-1500  
; TELEFAX: 703-415-1508  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-596-081A-14

Query Match 17.5%; Score 38.5; DB 1; Length 28;  
Best Local Similarity 38.5%; Pred. No. 1.2e+02;  
Matches 10; Conservative 4; Mismatches 7; Indels 5; Gaps 1;

QY 2 KLRFLQTEDSWVPAS-----PDTGLD 22  
| | | | | | | | | |  
Db 1 KKKFLNTEPSQLPPTAPPLPLPHSNLD 26

Search completed: January 26, 2005, 16:02:48  
Job time : 46 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 18:09:35 ; Search time 1821 Seconds  
(without alignments)  
7089.560 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_1\_273

Perfect score: 273  
Sequence: 1 atggagcactaccggaagc.....ccaagtaaccttcttcag 273

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 segs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 3273528

Minimum DB seq length: 0  
Maximum DB seq length: 273

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_cm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29.2	10.7	166	11	BV165390
2	28.8	10.5	133	4	AB126591 Canis fam
3	28.8	10.5	201	11	BV200943 sqnm20546
4	28.8	10.5	227	9	HS066605 Homo sapi
5	28.4	10.4	282	6	CQ660987 Sequence
6	28.2	10.3	273	11	MMSN122
7	28	10.3	273	10	L35014 Mouse N-met
8	27.8	10.2	154	6	CQ679231 Sequence
9	27.8	10.2	194	6	CQ443803 Sequence
10	27.4	10.0	240	9	HS0230371 Homo sapi
11	27.2	10.0	99	6	CQ734281 Sequence
12	27.2	10.0	235	6	AX798801 Sequence
13	27	9.9	160	11	RICG332A
14	27	9.9	241	9	HS140C7F
15	27	9.9	264	6	BD237006 Compounds
16	27	9.9	264	6	AR225406 Sequence
17	27	9.9	284	6	AX321476 Sequence
18	27	9.9	264	9	S68587 platelet-ty
19	27	9.9	264	9	S68588 platelet-ty

C	20	26.8	9.8	208	11	HS0310747	AJ310747 Homo sapi
	21	26.8	9.8	241	11	HSPE07C12	AL009365 H.sapiens
C	22	26.8	9.8	251	11	BV167732	BV167732 sqnm6347
C	23	26.8	9.8	269	9	HS035828	Z22256 H.sapiens D
C	24	26.6	9.7	245	5	AB035828	AB035828 Coturnix
	25	26.4	9.7	167	6	CQ080407	CQ080407 Sequence
	26	26.4	9.7	167	6	CQ114259	CQ114259 Sequence
	27	26.4	9.7	167	6	CQ153138	CQ153138 Sequence
	28	26.4	9.7	167	6	CQ186132	CQ186132 Sequence
	29	26.4	9.7	167	6	CQ236438	CQ236438 Sequence
	30	26.4	9.7	167	6	CQ274027	CQ274027 Sequence
	31	26.4	9.7	167	6	CQ311246	CQ311246 Sequence
	32	26.4	9.7	167	6	CQ348333	CQ348333 Sequence
C	33	26.4	9.7	198	6	CQ081138	CQ081138 Sequence
C	34	26.4	9.7	198	6	CQ115663	CQ115663 Sequence
C	35	26.4	9.7	198	6	CQ154439	CQ154439 Sequence
C	36	26.4	9.7	198	6	CQ186999	CQ186999 Sequence
C	37	26.4	9.7	198	6	CQ237699	CQ237699 Sequence
C	38	26.4	9.7	198	6	CQ275317	CQ275317 Sequence
C	39	26.4	9.7	198	6	CQ312286	CQ312286 Sequence
C	40	26.4	9.7	198	6	CQ349715	CQ349715 Sequence
C	41	26.2	9.6	240	9	HS0415844	AJ415844 Homo sapi
C	42	26	9.5	72	10	AF093879	AF093879 Mus muscu
C	43	26	9.5	92	6	AR204954	AR204954 Sequence
C	44	26	9.5	166	11	BV165391	BV165391 MsU63-13
C	45	26	9.5	171	3	AF135376	AF135376 Semibalan

ALIGNMENTS

RESULT 1	BV165390	166 bp	DNA	linear	STS 02-AUG-2004
LOCUS	MsU63-12	PCR fragment of the molecular marker, MsU63-Vr-TC1966			
DEFINITION	Vigna radiata STS genomic, sequence tagged site.				
ACCESSION	BV165390				
VERSION	BV165390.1	GI:50880891			
KEYWORDS	STS.				
SOURCE	Vigna radiata				
ORGANISM	Vigna radiata				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.				
REFERENCE	1 (bases 1 to 166)				
AUTHORS	Choi,H.K., Doyle,J. and Cook,D.R.				
TITLE	Development of Nuclear gene-derived Molecular Phylogenetic Markers Linked to Legume Genetic Maps				
JOURNAL	Unpublished (2004)				
COMMENT	Contact: Douglas R. Cook CAES Genomics Facility University of California-Davis, Plant Pathology One Shields Ave, Davis, CA 95616, USA Tel: 530 754 6561 Fax: 530 754 6617 Email: drcook@ucdavis.edu Primer A: GGACTTTCGTGCTACCCAGAGACA Primer B: AGCAAGGATGCTTTTGTGTCATGGA STS size: 166 PCR Profile: Presoak: 95 degrees C for 15.00 minutes Denaturation: 94 degrees C for 20 seconds Annealing: 55 degrees C for 20 seconds Polymerization: 72 degrees C for 1.00 minutes PCR Cycles: 35 Thermal cycler: MJ Research, Tetrad  Protocol: Template: 5-30 ng Primer: each 0.5 uM dNTP: each 250 uM Tag Polymerase: 0.05 units/ul				

```

Total Vol: 10 ul
Buffer:
  MgCl2 2.0 mM
  KCl 50 mM
  Tris-HCl 10 mM
  pH 8.7.
FEATURES
  source
    Location/Qualifiers
      1..166
        /organism="Vigna radiata"
        /mol_type="genomic DNA"
        /cultivar="TC1966"
        /db_xref="taxon:157791"
        /sex="Hermaphrodite"
        /tissue type="Leaf"
        /clone_lib="PCR fragment of the molecular marker,
        MSU63-Vr-TC1966"
        /note="PCR-amplified genomic fragment using cross-species
        gene-specific primer sets"
      1..166
        /gene="MsU63"
        /notes="Light harvesting chlorophyll a/b-binding (CAB)
        protein"
      <1..>166
        /gene="MsU63"
  ORIGIN
    Query Match 10.7%; Score 29.2; DB 11; Length 166;
    Best Local Similarity 50.0%; Pred. No. 1.5e+03;
    Matches 73; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
  QY 119 TGGGGTTGGCTCGGGTGGTGGAGGGCGGCGAGTCTCGGCATGTAGTCTTCAGGTT 178
  Db 21 TGGGGTTGGTCTCCCTGAAGTGTGGTCCGGACGGTGTGAAGTTCGGGGAGCGAGTGT 80
  QY 179 CTGCGAGGCTGCAGGAAGAGCTGTACGTGCCTGAGATGTCAAGCGCGGGTCCCGAG 238
  Db 81 GGTTCAGGCTGGGTCTCAGATATTCAGTACGGTGGGCTGTCTACTTGGGCAACCCAA 140
  QY 239 GCTGCACAGCTCACCAGCTACGT 264
  Db 141 GCCTGATCCATGCCACAAAGCATCCT 166
RESULT 2
AB126591/c
LOCUS 133 bp DNA linear MAM 20-MAY-2004
DEFINITION Canis familiaris DRD4 gene, intron 2, complete sequence, isolate:Q.
ACCESSION AB126591
VERSION AB126591.1 GI:47522547
KEYWORDS
SOURCE
ORGANISM
  Canis familiaris (dog)
  Canis familiaris
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
  1 Nara.H. and Inoue-Murayama,M.
  Novel Polymorphism of the canine dopamine receptor D4 gene intron
  II region
JOURNAL
  Unpublished
REFERENCE
  2 (bases 1 to 133)
  Nara,H. and Inoue-Murayama,M.
  Direct Submission
  Submitted (19-NOV-2003) Miho Inoue-Murayama, Gifu-University,
  Faculty of Agriculture; Yanagido 1-1, Gifu, Gifu 501-1193, Japan
  (E-mail:miho-40cc.gifu-u.ac.jp, Tel:81-58-293-2874 (ex.2874),
  Fax:81-58-293-2874)
FEATURES
  source
    Location/Qualifiers
      1..133
        /organism="Canis familiaris"
        /mol_type="genomic DNA"
        /isolate="Q"
        /db_xref="taxon:9615"
  ORIGIN
    Query Match 10.5%; Score 28.8; DB 11; Length 201;
    Best Local Similarity 60.0%; Pred. No. 1.9e+03;
    Matches 48; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
  QY 72 TACCTTTCAGATCGGGTTCGAGATGCAGCAAAATTCGCAACCTGTCGGGTGGCTCT 131
  Db 200 TAGGCTTCAGACGACGACCTTGGAGGCGAGCAGCAACCATCCGCTTTTGTGCT 141
  QY 132 GGGTCGGTTGGAGGGCGCA 151
  Db 140 AGGCGGTGGATGCCGTCA 121
RESULT 4

```

```

gene
  1..133
    /gene="DRD4"
intron
  1..133
    /gene="DRD4"
    /note="dopamine receptor D4"
    /number=2
ORIGIN
  Query Match 10.5%; Score 28.8; DB 4; Length 133;
  Best Local Similarity 54.8%; Pred. No. 1.9e+03;
  Matches 57; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
  QY 137 GGTGGAGGCGCGCAGTCTCGCATGTAGTGTTCAGGTTCTGCGAGGGCTGCAGGAA 196
  Db 116 GGGCGCGCGCGCGCGCGGCGGCGCTGCGAGGGCGGCGGCGGCGGCGGCGGCGGCG 57
  QY 197 AGGCTGTCTAGCTGCGCTGAGATTGTCAAGCGCGGGTCCCGGC 240
  Db 56 GGGCGGCTCGCGAGGCGGGGCGCTGCGAGGGCGGCGGCGGCGGCGGCGGCGGCGG 13
RESULT 3
BV200943/c
LOCUS 201 bp DNA linear STS 10-JUN-2004
DEFINITION sqm205467 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
tagged site.
ACCESSION BV200943
VERSION BV200943.1 GI:48169634
KEYWORDS
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 201)
  Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
  Cantor,C.R. and Braun,A.
  Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
  Regions
JOURNAL
  Genome Res. (2004) In press
COMMENT
  Contact: Andreas Braun
  Pharmaceuticals division
  Sequenom, Inc.
  3595 John Hopkins Court, San Diego, CA 92121, USA
  Tel: 18582029018
  Fax: 18582029020
  Email: abraun@sequenom.com
  Primer A: No primer sequence submitted
  Primer B: No primer sequence submitted
  STS size: 201.
FEATURES
  source
    Location/Qualifiers
      1..201
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /clone_lib="Human DNA (Sequenom)"
        <1..>201
  STS
  ORIGIN
    Query Match 10.5%; Score 28.8; DB 11; Length 201;
    Best Local Similarity 60.0%; Pred. No. 1.9e+03;
    Matches 48; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
  QY 72 TACCTTTCAGATCGGGTTCGAGATGCAGCAAAATTCGCAACCTGTCGGGTGGCTCT 131
  Db 200 TAGGCTTCAGACGACGACCTTGGAGGCGAGCAGCAACCATCCGCTTTTGTGCT 141
  QY 132 GGGTCGGTTGGAGGGCGCA 151
  Db 140 AGGCGGTGGATGCCGTCA 121
RESULT 4

```





Db 200 CTTGAGAGGAGCTCAAGCGCGCAGCTGTGGAAGCAGCTGCGCGGCTTGATCTCTAG 141

QY 136 CGGTTGGAGGCGCGCAGTGTCTGGCATGT 164

Db 140 CCGGTATTATCTGACAGCTCTCGGCATGT 112

RESULT 7

MUSNR2C01

LOCUS 273 bp DNA linear ROD 23-JAN-1995

DEFINITION Mouse N-methyl-D-aspartate receptor subunit NR2C (NMDA2C) gene, exon 1.

ACCESSION L35014

VERSION L35014.1 GI:602738

KEYWORDS N-methyl-D-aspartate receptor subunit NR2C; NMDA receptor subunit NR2C.

SEGMENT 1 of 16

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 273)

AUTHORS Suchanek,B., Seeburg,P.H. and Sprengel,R.

TITLE Gene structure of the murine N-methyl D-aspartate receptor subunit NR2C

JOURNAL J. Biol. Chem. 270 (1), 41-44 (1995)

MEDLINE 95113860

PUBMED 7814402

COMMENT Original source text: Mus musculus (strain SV129) (tissue library: Stratagene Lambda FIX II) adult liver DNA.

FEATURES

source 1..273

1 /organism="Mus musculus"

2 /mol\_type="genomic DNA"

3 /strain="SV129"

4 /db\_xref="taxon:10090"

5 /tissue\_type="liver"

6 /dev\_stage="adult"

7 /tissue\_lib="Stratagene Lambda FIX II"

8 11..263

9 /gene="NMDA2C"

10 /note="5' untranslated exon; NR2C mouse gene; transcriptional start sites at basepairs 11, 12, 29"

11 /number=1

exon

ORIGIN

Query Match 10.3%; Score 28; DB 10; Length 273;

Best Local Similarity 56.5%; Pred. No. 3.3e+03;

Matches 52; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 90 CCGAGATGCGCAGCAAAATTCGCAACCTGCTGGGGTTGGCTCTGGGTGGTGGAGGGCG 149

Db 159 CCGAGCTCGCAGCCAGCGTCCAGCTCACGGGGCTGGCTTGGGACGACCGCGCG 218

QY 150 CAGTGTCTGGCATGTAGTTCTCAGGTTCTG 181

Db 219 GAGACGGCCCGGGCTCTGGGGTGGGTGCAG 250

RESULT 8

CO679231/c

LOCUS 154 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 24157 from Patent WO02070737.

ACCESSION CQ679231

VERSION CQ679231.1 GI:42191392

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1

AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.

TITLE Compositions and methods relating to osteoarthritis

JOURNAL Patent: WO 02070737-A 24157 12-SEP-2002;

Chondrogene Inc. (CA)

FEATURES

source Location/Qualifiers

1..154

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 10.2%; Score 27.8; DB 6; Length 154;

Best Local Similarity 56.2%; Pred. No. 3.7e+03;

Matches 50; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 63 ACTCTGTATACCTTGAGATGCGGGTCGAGATGCGCAAAATTCGCAACCTCTCTGG 122

Db 153 ACTTTCTTTAGGCTTCANACGCCGACCTTNAGGCGCAGGACCCACCATCCGCTTTT 94

QY 123 GTTGGCTCTGGGTGGTTGGAGGGCGCA 151

Db 93 TCTTGTCTAGGGCGGTGGATGCCGTCA 65

RESULT 9

CO443803

LOCUS 194 bp DNA linear PAT 30-JAN-2004

DEFINITION Sequence 9563 from Patent WO0192523.

ACCESSION CQ443803

VERSION CQ443803.1 GI:41416132

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1

AUTHORS Shimkets,R.A. and Leach,M.D.

TITLE Human polynucleotides and polypeptides encoded thereby

JOURNAL Patent: WO 0192523-A 9563 06-DEC-2001;

Curagen Corporation (US)

FEATURES

source Location/Qualifiers

1..194

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 10.2%; Score 27.8; DB 6; Length 194;

Best Local Similarity 59.5%; Pred. No. 3.7e+03;

Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 171 CTCAGGTTCTGCGAGGCTGCAGGAAGGCTGTCAGCTGCGCTGAGATTGTCAAGCGCG 230

Db 11 CTCAGGTCGGGGGTGGAGTAACGAGCAGATGAGCTGCCCTGAGATTGCTGGAGGAG 70

QY 231 GGTCCCGAGGCTGCACCCAG 249

Db 71 GGAAGCGGGCAGACCCAG 89

RESULT 10

HSA230371/c

LOCUS 240 bp DNA linear PRI 02-MAR-1999

DEFINITION Homo sapiens BF2N3-L3-C11 gene for immunoglobulin lambda chain variable region.

ACCESSION AJ230371

VERSION AJ230371.1 GI:4468513

KEYWORDS immunoglobulin; immunoglobulin lambda chain; immunoglobulin superfamily; joining region; variable region.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 240)

AUTHORS Farner,N.L.

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 240)  
AUTHORS Farnier, N.L.  
TITLE Direct Submission  
JOURNAL Submitted (05-AUG-1998) Farnier N.L., Department of Internal Medicine, Harold C. Simmons Arthritis Research Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235-8884, USA

FEATURES  
source  
1. .240  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="donor JW"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/cell\_line="single B lymphocyte"  
/cell\_type="CD5(-) B lymphocyte"  
/tissue\_type="blood"  
/rearranged  
1. .240  
/gene="BF2N3-L3-C11"  
1. .240  
/gene="BF2N3-L3-C11"  
/note="3J/23/P"

ORIGIN  
Query Match 10.0%; Score 27.4; DB 9; Length 240;  
Best Local Similarity 49.0%; Pred. No. 4.9e+03;  
Matches 73; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 103 AAATTGCGAACCTGCTGGGTTGGCTCTGGTGGTGGAGCGCGCAGTCTCGGCAT 162  
DB 201 ATAGTCAGCTCATCCCGCTTGGCTCTGCTGATGGTCAGGTGGCCGTGTTCCCGA 142

QY 163 GTAGTGTCTCAGGTTCTGCGAGGCTGCGAGGAAGCTGTCAGCTCGCTGAGATTGC 222  
DB 141 GTTGGAGCCAGAAATCGCTCAGGATCCAGAGGCCCGTTGTATCCCTATAGATGAC 82

QY 223 AAGCGCGGGTCCCGAGCCCTGCACCACT 251  
DB 81 CAGCACAGGGGCTGGACTGCTTCTGCT 53

RESULT 11  
CQ734281/c  
LOCUS CQ734281 99 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 20215 from Patent WO02068579.  
ACCESSION CQ734281  
VERSION CQ734281.1 GI:42321535  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof  
JOURNAL Patent: WO 02068579-A 20215 06-SEP-2002;  
PE Corporation (NY) (US)

FEATURES  
source  
1. .99  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 10.0%; Score 27.2; DB 6; Length 99;  
Best Local Similarity 67.9%; Pred. No. 5.5e+03;  
Matches 38; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 176 GTTCTGGCAGGCTGCAGGAAGCTGTCAGCTGCGCTGAGATTGTCAAGCGCGG 231

Db 59 GTTTATGAGGGCTGCAGTCAATGCTGACATGTGCTGAAGTCTTCAATCAGCTG 4

RESULT 12  
AX798801/c  
LOCUS AX798801 235 bp DNA linear PAT 08-OCT-2003  
DEFINITION Sequence 287 from Patent WO03054177.  
ACCESSION AX798801  
VERSION AX798801.1 GI:37604879  
KEYWORDS unidentified  
SOURCE unidentified  
ORGANISM unclassified.

REFERENCE 1  
AUTHORS Breves, R., Maurer, K.H., Eck, J., Lorenz, P. and Zinke, H.  
TITLE New glycosyl hydrolases  
JOURNAL Patent: WO 03054177-A 287 03-JUL-2003;  
Henkel Kommanditgesellschaft auf Aktien (DE)  
Location/Qualifiers  
1. .235  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="Metagenom"

ORIGIN  
Query Match 10.0%; Score 27.2; DB 6; Length 235;  
Best Local Similarity 50.0%; Pred. No. 5.6e+03;  
Matches 68; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 115 CTGCTGGGTTGGCTCTGGGTCGGTGGAGGGCGGAGTGTCTCGCATGTAGTGTCTCA 174  
DB 205 CTCCTTGGCCAGGTGACCGAGTCCGAGTCCGAGTCCCATCATGCGGTTCTCGT 146

QY 175 GGTTCGGCAGGCTTCAGGAAGGCTGTGAGTGTGAGTGTGCAAGCGCGGCTC 234  
DB 145 GCGGCTGTTCGAGCAGCAGGTGCTGTGGCCAGCGGTCGCGAGCGCGCTGGTA 86

QY 235 CCAGGCTGCACGAC 250  
DB 85 GTAGCCGGCAGCATC 70

RESULT 13  
RICG332A/c  
LOCUS RICG332A 160 bp DNA linear STS 23-MAR-2002  
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, G332A, sequence tagged site.  
ACCESSION D25339  
VERSION D25339.1 GI:436648  
KEYWORDS STS; Genomic; RFLP; Rice.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 160)  
AUTHORS Minobe, Y.  
TITLE Nucleotide sequence of sequence tagged site from japonica rice Nipponbare as an RFLP marker  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 160)  
AUTHORS Minobe, Y.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-1993) Yuza Minobe, National Institute of Agrobiological Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305, Japan  
COMMENT (E-mail: MINOBE@rics0.riken.go.jp, Tel: 0298-38-7441, Fax: 0298-38-7468)  
Submitted (18-Nov-1993) to DDBJ by:  
Yuza Minobe  
Dept. Rice Genome Research Program

## National Institute of Agrobiological Resources

Kannondai 2-1-2  
Tsukuba, Ibaraki  
JapanPhone: 0298-38-7441  
Fax: 0298-38-7468  
PROJECT = 'RGP'

## FEATURES

source

Location/Qualifiers

1..160  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"

## ORIGIN

## Query Match

Best Local Similarity 56.0%; Score 27; DB 11; Length 160;

Matches 51; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

## QY

130 CTGGTTCGGTGGAGGGCGGCGATGCTCGGATGATGTTCTCAGGTTCTGCGAGGCT 189

## Db

126 CTGGCTGTGAGCGCGCGGCGATGATGTTCTGCGGCTCTCTGGTGGCGTGA 67

## QY

190 GCAGAAAGGCTGTCAGCTGCGTGAGATTG 220

## Db

66 GCAGCAGTCGATGCTCTCGATGACGAGGATG 36

## RESULT 14

## HS140C7F

## LOCUS

HS140C7F 241 bp DNA linear PRI 19-OCT-1995  
H.sapiens Cpg island DNA genomic Mael fragment, clone 140c7,  
forward read cpg140c7.ft1a.

## ACCESSION

## Z59291

## VERSION

## Z59291.1

## KEYWORDS

## Cpg island; genomic Mael fragment.

## SOURCE

## Homo sapiens (human)

## ORGANISM

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## 1

## Cross, S.H., Charlton, J.A., Nan, X. and Bird, A.P.

## Purification of Cpg islands using a methylated DNA binding column

## Nat. Genet. 6 (3), 236-244 (1994)

## MEDLINE

## 94282070

## PUBMED

## 8012384

## REFERENCE

## 2

## (bases 1 to 241)

## Dodsworth, S.J., Huckle, E., Wilkinson, P. and Micklem, G.

## Direct Submission

## Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,

## CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk

## Vector: pGEM-52f(-)

## Clones are available from the UK MRC Human Genome Mapping Project

## Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:

## http://www.hgmp.mrc.ac.uk/ for details

## or contact: biohelp@hgmp.mrc.ac.uk.

## Location/Qualifiers

## 1..241

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## /clone="140c7"

## /sex="male"

## /tissue\_type="blood"

## /clone\_lib="CGI-1"

## /dev\_stage="adult"

## ORIGIN

## Query Match

Best Local Similarity 51.7%; Score 27; DB 9; Length 241;

Matches 60; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

## QY

60 GCTACCTCTGTATACCTTGTAGATGCGGGTCCGAGATGCGAGAAATTCGCAACCTGCT 119

Db 122 GCTGCCACTGTGGCGCGGAGAGCGGCTCTGTGNCAGATAGATCCCTAGATTCA 181  
QY 120 GGGGTTGGCTCTGGGTCTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCTCAG 175  
Db 182 GGCCCGGCTGGGCGCTGGAGGGACCGCTGGGCTTCGTAGTAGATTCTCCG 237

## RESULT 15

## BD237006

## LOCUS

## DEFINITION

## Compounds for remedy and diagnosis of lung cancer and method for

## using the same.

## ACCESSION

## BD237006

## VERSION

## BD237006.1

## KEYWORDS

## JP 2002516659-A/7.

## SOURCE

## Homo sapiens (human)

## ORGANISM

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## 1

## (bases 1 to 264)

## Reed, S.G., Lodes, M.J., Frudakis, T.N. and Mohamath, R.

## Compounds for remedy and diagnosis of lung cancer and method for

## using the same

## TITLE

## Patent: JP 2002516659-A 7 11-JUN-2002;

## JOURNAL

## CORIXA CORP

## COMMENT

## OS Homo sapiens (human)

## PN JP 2002516659-A/7

## PD 11-JUN-2002

## PF 26-JAN-1999

## JP 2000529432

## PR 28-JAN-1998

## US 09/015029

## 28-JAN-1998

## US 09/040828

## 18-MAR-1998

## US 09/040828

## 23-JUL-1998

## US 09/122192

## 22-DEC-1998

## US 09/219245

## PI STEVEN G REED, MICHAEL J LODES, TONY N FRUDAKIS, RAODOH MOHAMATH

## PC C12N15/09, A61K35/14, A61K38/00, A61K39/00, A61K39/39, A61K39/395,

## PC A61K39/395,

## PC A61P11/00, A61P35/00, C07K14/47, C07K16/18, C07K19/00, C12N1/19, PC

## C12N1/21,

## PC C12N5/10, C12P21/08, C12Q1/68, G01N33/53, G01N33/574, G01N33/577//

## PC (C12N1/21, C12R1:19), C12N15/00, A61K37/02, C12N5/00 CC n =

## A, T, C or G

## FH Key

## misc\_feature (1)..(264).

## FT Location/Qualifiers

## 1..264

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## ORIGIN

## Query Match

Best Local Similarity 54.3%; Score 27; DB 6; Length 264;

Matches 51; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

## QY

175 GGTCTGCGAGGCTGCGAGGAAAGCTCTCAGCTGCGGTGAGATTCTCAAGCGCGGCTC 234

## Db

12 GATGGGCGACGGGAGACAGATGACCTTTAACTGCCCCACGTTNTCMAGGAAGGATT 71

## QY

235 CCAGGCTGCGACCGAGCTCACCAGCTACGTTTCC 268

## Db

72 ACAGCGGTGAGCCACTGCGCGCGGCTTTCTCC 105

Search completed: January 31, 2005, 19:44:40  
Job time : 1826 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 18:58:47 ; Search time 1938 Seconds  
(without alignments)  
5133.151 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_1\_273

Perfect score: 273  
Sequence: 1 atggagcatatccggaagc.....ccaagtacgtttcttcag 273

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 7660530

Minimum DB seq length: 0  
Maximum DB seq length: 273

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gssi:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	225.4	82.6	272	4	BM699759 UI-E-DWI-
2	206.2	75.5	268	1	AA022458 ze70h07.r
3	138.4	50.7	194	2	AW436738 77186 MAR
4	135	49.5	220	6	CB117372 K-EST0162
5	117.4	43.0	231	7	W48348 mc86b02.r1
6	103	37.2	256	1	AI183131 ub93g04.r
7	101.6	37.2	217	1	AA239107 mv36c09.r
8	97.6	35.8	246	7	W07888 mb49d03.r1
9	96.4	35.3	188	1	AA756051 vul1e10.r
10	78.2	28.6	120	1	AI391194 mb49d03.y
11	46.6	17.1	263	2	AW326399 18938 MAR
12	33.8	12.4	247	4	BM068792 1e85a09.y
13	33	12.1	230	1	AV435580
14	32.4	11.9	220	2	AW818811 RC4-ST027
15	32.2	11.8	133	1	AV430042
16	32.2	11.8	181	1	AV432681
17	32.2	11.8	204	1	AV434181
18	32.2	11.8	206	1	AV435325
19	32.2	11.8	210	1	AV432576
20	32.2	11.8	224	1	AV433480
21	32.2	11.8	237	1	AV436270
22	31.6	11.6	231	2	BF478851
23	31.4	11.5	240	2	BE774505
24	31.4	11.5	250	2	BE771297 QV1-FT008

25	31.4	11.5	258	6	CB885146
26	31.4	11.5	262	4	BG953905
27	31.2	11.4	268	2	BE937539
28	31	11.4	247	2	BE717545
29	31	11.4	260	7	CN363918
30	30.8	11.3	233	1	AV435417
31	30.8	11.3	246	1	AV435673
32	30.8	11.3	249	1	AV432026
33	30.8	11.3	266	9	FR0044635
34	30.6	11.2	135	1	AV430758
35	30.6	11.2	229	4	BJ162793
36	30.4	11.1	121	1	AV435623
37	30.4	11.1	188	1	AL045467
38	30.4	11.1	223	7	CO264753
39	30.4	11.1	224	1	AA301864
40	30.4	11.1	230	1	AV435742
41	30.4	11.1	236	4	BM740588
42	30.4	11.1	241	1	AV435395
43	30.2	11.1	232	1	AI508914
44	30.2	11.1	251	7	R25661
45	30.2	11.1	252	7	CO698646

#### ALIGNMENTS

RESULT 1  
BM699759  
LOCUS BM699759 272 bp mRNA linear EST 28-FEB-2002  
DEFINITION UI-E-DWI-agz-i-16-0-UI.r1 UI-E-DWI Homo sapiens cDNA clone  
UI-E-DWI-agz-i-16-0-UI 5', mRNA sequence.  
ACCESSION BM699759  
VERSION GI:19013017  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 272)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE Ronaldo, M.F., Lennon, G. and Soares, M.B.  
Normalisation and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1..272  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-DWI-agz-i-16-0-UI"  
/issue\_type="lens"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-DWI"  
/note="Organ: eye; Vector: p77T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-DWI is a normalized cDNA library containing the following tissue(s): lens. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

Query Match 82.6%; Score 225.4; DB 4; Length 272;  
Best Local Similarity 99.1%; Pred. No. 1.5e-51;  
Matches 226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGGACATACCGGAAAGCTGCTGTGTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 60  
DB 45 ATTGAGGACATACCGGAAAGCTGCTGTGTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 104

QY 61 CTACCTCTGTATACCTTGTAGATCGGGTCCGAGATCGAGCAAAATTCGCACTGCTG 120  
DB 105 CTACCTCTGTATACCTTGTAGATCGGGTCCGAGATCGAGCAAAATTCGCACTGCTG 164

QY 121 GGGTTGGCTCTGGGTGCGGTTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
DB 165 GGGTTGGCTCTGGGTGCGGTTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 224

QY 181 GCGAGGCTGCGAGAAAGCTGTGAGTGGCGGTGAGATGTCGAGCGG 228  
DB 225 GCGAGGCTGCGAGAAAGCTGTGAGTGGCGGTGAGATGTCGAGCGG 272

## RESULT 2

AA022458 268 bp mRNA linear EST 30-JAN-1997  
LOCUS ze70h07.r1 Soares fetal heart NBH19W Homo sapiens cDNA clone  
DEFINITION IMAGE:364381 5', mRNA sequence.  
ACCESSION AA022458  
VERSION AA022458.1 GI:1486557  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 268)  
AUTHORS Hillier, J., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissole, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
MEDLINE 97044478  
PUBMED 889549  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL, contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 945 Std Error: 0.00  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 200.  
Location/Qualifiers  
1. .268  
/organism="Homo sapiens"  
/mol\_type="mRNA"

FEATURES  
source

/db\_xref="GDB:1281085"  
/db\_xref="taxon:9606"  
/clone="IMAGE:364381"  
/sex="unknown"  
/dev\_stage="19 weeks"  
/lab\_host="PH108 (ampicillin resistant)"  
/clone\_lib="Soares fetal heart NBH19W"  
/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBH19W."

## ORIGIN

Query Match 75.5%; Score 206.2; DB 1; Length 268;  
Best Local Similarity 97.7%; Pred. No. 3.1e-46;  
Matches 208; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 60  
DB 54 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 113

QY 61 CTACCTCTGTATACCTTGTAGATCGGGTCCGAGATGCGAGCAAAATTCGCACTGCTG 120  
DB 114 CTACCTCTGTATACCTTGTAGATGCGGGTCCGAGATGCGCAAAATTCGCACTGCTG 173

QY 121 GGGTTGGCTCTGGGTGCGGTTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
DB 174 GGGTTGGCTCTGGGTGCGGTTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 233

QY 181 GCGAGGCTGCGAGAAAGCTGTGAGTGGCGGTGAGATGTCGAGCGG 213  
DB 234 GCGAGGCTGCGAGAAAGCTGTGAGTGGCGGTGAGATGTCGAGCGG 266

## RESULT 3

AA0436738 194 bp mRNA linear EST 09-JUL-2000  
LOCUS AW436738  
DEFINITION 77186 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION AW436738  
VERSION AW436738.1 GI:6972044  
KEYWORDS EST.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa  
REFERENCE 1 (bases 1 to 194)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus. Fahrenkrug, S.C., Smith, T.P.L., Preking, B.A., Cho, J., White, J., Vallet, J., Wise, T., Rohrer, G.A., Pettea, G., Sultana, R., Quackenbush, J., and Keeler, J.W.

TITLE Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly  
JOURNAL Mamm. Genome 13 (8), 475-478 (2002)  
MEDLINE 22213789  
PUBMED 12226715  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 20 and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTCCTCCAGTCACGACG  
Plate: 32 row: M column: 15  
Seq primer: ATTAGGTGACACTATAG.

## FEATURES

Location/Qualifiers

source

1..194

/organism="Sus scrofa"

/mol\_type="mRNA"

/db\_xref="taxon:9823"

/tissue\_type="pooled"

/lab\_host="DH10B"

/clone\_lib="WARC-2PIG"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."

## ORIGIN

Query Match 50.7%; Score 138.4; DB 2; Length 194;  
Best Local Similarity 90.2%; Pred. No. 1.6e-27;  
Matches 148; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 30 ATGGAGCACTACCGGAAAGCTGGCTCTGTGGAATACCAGCACTTCCCAATGCCCCAG 89  
QY 61 CTACCTCTCATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 90 CTGCTCTCATACCTTGGAGATCGGAGTCCGAGATGGCAGCAAGATCCGCAACCTGCTG 149  
QY 121 GGGTTGGCTCTGGTGGTGGAGGGCGGAGTGGCTGGCATGT 164  
Db 150 GGACTGGCTGGTGGGCGGTTGGAGGGTGGCGAGCGCCCGGCATGT 193

## RESULT 4

CB117372 220 bp mRNA linear EST 28-JAN-2003  
LOCUS K-EST0162559 L4SNJ368 Homo sapiens CDNA clone L4SNJ368-30-H04 5',  
mRNA sequence.  
CB117372  
CB117372.1 GI:27943179  
EST.  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 220)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

21C Frontier Korean EST Project 2001

## TITLE

## JOURNAL

## COMMENT

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 30 row: H column: 04

High quality sequence stop: 220.

Location/Qualifiers

source

1..220

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="L4SNJ368-30-H04"

/sex="M"

/tissue\_type="Liver"

/cell\_type="Polygonal"

/cell\_line="SNU-368"

/lab\_host="Top10F"

/clone\_lib="L4SNJ368"

/note="Organ: Liver; Vector: pcNS-D2; Site\_1: EcoRI;

Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including  
EcoRI site by treatment of T4 RNA ligase and the first  
strand cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The cDNA vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F' by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

## ORIGIN

Query Match 49.5%; Score 135; DB 6; Length 220;  
Best Local Similarity 100.0%; Pred. No. 1.4e-26;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 86 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 145  
QY 61 CTACCTCTCTGATACCTTGCAGATGGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 146 CTACCTCTCTGATACCTTGCAGATGGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 205  
QY 121 GGGTTGGCTCTGGGT 135  
Db 206 GGGTTGGCTCTGGGT 220

## RESULT 5

W48348 251 bp mRNA linear EST 24-MAY-1996  
LOCUS mc86b02.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus CDNA  
clone IMAGE:355371 5', mRNA sequence.  
W48348  
W48348.1 GI:1333399  
EST.  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 251)  
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisell,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.

## TITLE

## JOURNAL

## COMMENT

Unpublished (1996)

Contact: Marra M/Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:227171

Seq primer: mob.REGA+ET

High quality sequence stop: 201.

Location/Qualifiers

source

1..251

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="IMAGE:355371"

```

/db_xref="taxon:10090"
/clone="IMAGE:1396086"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NBMWG"
/notes="Organ: mammary gland; vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGCAATTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2 ]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

```

## ORIGIN

```

Query Match 43.0%; Score 117.4; DB 7; Length 251;
Best Local Similarity 82.1%; Pred. No. 1e-21;
Matches 147; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 1 ATGAGACACTACCGAAAGTGGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 60
    |||||
Db 74 ATGGAACAGTACCGG-AGGCGGCTCTGTAGAGCTTCAGCCTCATCACCAATGCCCGAG 132
    |||||

QY 61 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATCCCAACCTGCTG 120
    |||||
Db 133 CTACCTCTGACACTCTGGAATCGGATCGGATGGAGCAAAATCCCAACCTGCTA 192
    |||||

QY 121 GGGTGGCTCTGGTCTGGTGGAGCGCGAGTCTCGCATGTAGTGTCTCAGGTTTC 179
    |||||
Db 193 GGGCTGGCGTGGTCTGGTGGAGCGGAGGAGCAGGACGCTGGTGTCTCAGGCTC 251
    |||||

```

## RESULT 6

```

A1183131
LOCUS 256 bp mRNA linear EST 08-OCT-1998
DEFINITION ub93G04.r1 Soares mammary_gland_NBMWG Mus musculus cDNA clone
IMAGE:1396086 5', mRNA sequence.
A1183131
VERSION A1183131.1 GI:3733769
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 256)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:907802
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 252.
Location/Qualifiers
1. .256
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

```

## FEATURES

```

source
Location/Qualifiers
1. .256
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

```

```

/db_xref="taxon:10090"
/clone="IMAGE:1396086"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NBMWG"
/notes="Organ: mammary gland; vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGCAATTTTTTTTTTTTTTTT T 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

```

## ORIGIN

```

Query Match 37.7%; Score 103; DB 1; Length 256;
Best Local Similarity 81.7%; Pred. No. 9.9e-18;
Matches 143; Conservative 0; Mismatches 30; Indels 2; Gaps 2;

QY 1 ATGAGACACTACCGAAAGTGGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 60
    |||||
Db 84 ATGGAACAGTACCGG-AGGCGGCTCTGTAGA-CTTCAGCCTCATCACCAATGCCCGAG 141
    |||||

QY 61 CTACCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 120
    |||||
Db 142 CTACCTCTGACACTCTGGAATCGGATCGGATGGAGCAAAATTCGCAACCTGCTA 201
    |||||

QY 121 GGGTGGCTCTGGTCTGGTGGAGCGCGAGTCTCGGCATGTAGTGTCTCAG 175
    |||||
Db 202 GGGCTGGCGTGGTCTGGTGGAGCGGAGCAGCAGGCGTGTGTCTCAG 256
    |||||

```

## RESULT 7

```

A18239107
LOCUS 217 bp mRNA linear EST 03-MAR-1997
DEFINITION my36C09.r1 Barstead mouse pooled organs MPLRB4 Mus musculus CDNA
clone IMAGE:697936 5', mRNA sequence.
A18239107
VERSION A18239107.1 GI:1863130
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 217)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:431496
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 169.
Location/Qualifiers
1. .217
/organism="Mus musculus"
/mol_type="mRNA"

```



```

/dev_stage="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:697936"
/sex="mixed"
/tissue_type="pooled organs"
/dev_stage="7 day"
/lab_host="DH10B"
/clone_lib="Barstead mouse pooled organs MPLRB4"
/notes="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTAGATCTGAAGTGGAGCGGCGCCCTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [GTGATTGGTACC], digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library constructed by Bob Barstead."

ORIGIN

Query Match      37.2%; Score 101.6; DB 1; Length 217;
Best Local Similarity 83.6%; Pred. No. 2.4e-17;
Matches 127; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 1 ATGGAGCACTACCGGAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCCGAG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 ATGGAACAGTACCGG-AGGCCGCTCTGTAGAGCTTCCAGCCTCATCAATGCCCGAG 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 CTACCTCTGATACCTCTGATGCGGGTCCGAGATGGCGACAAATTCGCAACCTGCTG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 CTACCTCTGACACTCTGGAATGCGAGTCCGAGTGGGAGCAATTCGCAACCTGCTA 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GGGTTGGCTCGGTGGTGGAGGGCGGAG 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 GGGCTGGCGTGGTGGTGGTGGAGGGGGAAG 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 8
W07888
LOCUS
DEFINITION
  mb49403.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:332741 5', mRNA sequence.
ACCESSION
  W07888
VERSION
  W07888.1 GI:1281991
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 246)
REFERENCE
  AUTHORS
    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
    Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
    Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
    Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
    Waterston,R.
  The WashU-HMI Mouse EST Project
  Unpublished (1996)
  CONTACT: Marra M/Mouse EST Project
  WashU-HMI Mouse EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LLNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  MGI:214141
  Seq primer: mob.REGA+ET
  High quality sequence stop: 237.
  Location/Qualifiers
    1..246
    /organism="Mus musculus"
    /mol_type="mRNA"
    /db_xref="taxon:10090"
    /clone="IMAGE:332741"

FEATURES
  source

```

```

/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/notes="Vector: p7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCGCCATTTTTTTTTTTTTT 3']; double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaudo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match      35.8%; Score 97.6; DB 7; Length 246;
Best Local Similarity 83.1%; Pred. No. 3.1e-16;
Matches 123; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 125 TGCTCTCGGTGGTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTCTCGGCA 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 TGGCGTGGGTGTTTGGAGGGGAGCAGGACGCTGTCTCAGGTCTCGGC- 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 GGGCTCAGAAAGGCTGTGAGTGGCTGAGATTCTCAAGCGGCGGTCCTCCAGGCTGC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GTGCTGCTGGAAAGGCTGTGAGTGTGAGAGATTGTCAAAAGGCGGTTCCGGGCTAC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 245 ACCAGCTCACCAGCTACGTTTCTTCA 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 ACCAGCTCACCAGCTCCTCTCTGCA 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 9
AA756051
LOCUS
DEFINITION
  vulle10.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
IMAGE:1180362 5', mRNA sequence.
ACCESSION
  AA756051
VERSION
  AA756051.1 GI:2803249
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 188)
REFERENCE
  AUTHORS
    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
    Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
    Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
    Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
    Waterston,R.
  The WashU-HMI Mouse EST Project
  Unpublished (1996)
  CONTACT: Marra M/Mouse EST Project
  WashU-HMI Mouse EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LLNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  MGI:638210
  Seq primer: -28ml3 rev2 ET from Amersham
  High quality sequence stop: 104.
  Location/Qualifiers
    1..188
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C3H"
    /db_xref="taxon:10090"
    /clone="IMAGE:1180362"
    /cell_line="C2C12"

FEATURES
  source

```



/note="vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."

```

ORIGIN
Query Match      17.1%; Score 46.6; DB 2; Length 263;
Best Local Similarity 92.5%; Pred. No. 0.038; 4; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 0;

QY 1 ATGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAAT 53
|||||
Db 211 ATGAGCACTACCGGAAGCTGGCTCTGTGGAATCACCACACTTCCCAAT 263
|||||

RESULT 12
BM068792/c
LOCUS
DEFINITION
cDNA clone IMAGE:5673520 5', mRNA sequence.
ACCESSION
BM068792
VERSION
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Melton D., Brown J., Kenty G., Permutt A., Lee C., Kaestner K.,
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 247)
AUTHORS
Lemishka, I., Secare, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V.,
Williams, T., Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
TITLE
Unpublished (2000)
JOURNAL
Other ESTs: ie85a09.xl
COMMENT
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biochem.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 221.
FEATURES
Location/Qualifiers
1..247
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5673520"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
Site 2: Sal I; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated

```

by hydroxyapatite chromatography and used to make this  
library."

```

ORIGIN
Query Match      12.4%; Score 33.8; DB 4; Length 247;
Best Local Similarity 60.2%; Pred. No. 1.3e+02;
Matches 56; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 36 CCAGGCGCTTCCCAATGCCAGCTACTCTCTGATACCTTGAGATGGGCTCCGAGA 95
|||||
Db 113 CCAGGCTCGGCGCTTCCAGACCTGGCGCGGGGTCACATCGGGTCAAGGA 54
|||||

QY 96 TGCGACAAATTCGCAACCTCTGGGGTTGGC 128
|||||
Db 53 AGGCAGCAAGATCCGTAACCTGATGCGCTTCGC 21
|||||

RESULT 13
AV435580/c
LOCUS
DEFINITION
230 bp mRNA linear EST 23-AUG-2000
AV435580 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PM059e02 r 5', mRNA sequence.
ACCESSION
AV435580
VERSION
GI:8590805
SOURCE
Porphyra yezoensis
ORGANISM
Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
1 (bases 1 to 230)
AUTHORS
Nikaido, I., Asamizu, E., Nakajima, M., Nakamura, Y., Saga, N. and
Tabata, S.
Generation of 10,154 expressed sequence tags from a leafy
gametophyte of a marine red alga, Porphyra yezoensis
DNA Res. 7, 223-227 (2000)
JOURNAL
MEDLINE
20363100
PUBMED
10907854
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 252-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
1..230
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/db_xref="taxon:2798"
/clone_lib="PM059e02 r"
/clone_lib="Porphyra yezoensis TU-1"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
ORIGIN
Query Match      12.1%; Score 33; DB 1; Length 230;
Best Local Similarity 54.5%; Pred. No. 2.1e+02;
Matches 66; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 95 ATGGCAGCAAAATTCGCAACCTGCTGGGGTTGGCTCTGGTTCGGTGGGCGGAGTG 154
|||||
Db 121 ATGACGCGCACCTGCGGGCGGGGCTAGGACGGGAGCCCGCGGGCTGG 62
|||||

QY 155 CTCGGCATGTAGTGTCTCAGGTTCTGGCAGGGCTGCAGGAAGGCTGTACAGTGGCGCTG 214
|||||
Db 61 CTGGGGATGGCGGGGCTGGGCAGGTGGTGGCGGGTGGATAGGGGGGAGGTGAGCTC 2
|||||

QY 215 A 215
1 A 1
RESULT 14
AW818811/c

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 18:07:55 ; Search time 317 Seconds  
(without alignments)

4520.795 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_1\_273

Perfect score: 273

Sequence: 1 atggagcactacggaagc.....ccaagtaagttcttcctcg 273

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 4923180

Minimum DB seq length: 0

Maximum DB seq length: 273

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	29.8	10.9	192	6	ABQ55864 Human ova
C 2	28.6	10.5	265	12	ACH82956 Human gen
C 3	27.8	10.2	194	6	ABN20543 Human ORF
C 4	27.6	10.1	122	12	ACH85301 Human gen
C 5	27.2	10.0	133	12	ADQ091957 Dog dopam
C 6	27.2	10.0	235	10	ADF31067 Soil meta
7	27	9.9	264	3	AAC79052 Human lun
8	27	9.9	264	4	AAD23127 Human lun
9	27	9.9	264	10	ADD67064 Human lun
10	27	9.9	264	10	ADB87569 Human lun
11	26.4	9.7	167	4	AAI26274 Probe #16
12	26.4	9.7	167	4	ABA73983 Human foe
13	26.4	9.7	167	4	AAI54432 Probe #23
14	26.4	9.7	167	4	ABA39062 Probe #17
15	26.4	9.7	167	4	Aak48603 Human bron
16	26.4	9.7	167	4	Aak22436 Human bra
17	26.4	9.7	167	4	ABS48287 Human liv
18	26.4	9.7	167	6	ABS22360 Human gen
C 19	26.4	9.7	198	4	AAI27005 Probe #16
C 20	26.4	9.7	198	4	ABA75273 Human foe
C 21	26.4	9.7	198	4	AAI55836 Probe #24

C 22	26.4	9.7	198	4	ABA39929 Probe #18
C 23	26.4	9.7	198	4	Aak49904 Human bon
C 24	26.4	9.7	198	4	Aak23818 Human bra
C 25	26.4	9.7	198	4	ABs49548 Human liv
C 26	26.4	9.7	198	6	ABs23400 Human gen
C 27	26.2	9.6	216	3	AAF10329 FUBarium
C 28	26	9.5	198	4	ABA41025 Probe #19
C 29	26	9.5	249	8	ABT14594 Pseudomon
C 30	25.8	9.5	165	2	AAX51903 Human sec
C 31	25.8	9.5	169	6	ABL80211 Human ova
C 32	25.8	9.5	187	10	ACD94768 Human col
C 33	25.6	9.4	87	6	ADH31736 Yeast sma
C 34	25.6	9.4	212	10	ABX83988 Corn ear-
C 35	25.6	9.4	213	12	ACH86002 Human gen
C 36	25.6	9.4	261	11	ACH95344 Klebeie11
C 37	25.4	9.3	194	3	AAC18105 Human sec
C 38	25.4	9.3	235	10	ACA56063 Human sig
C 39	25.4	9.3	235	12	AD155859 Human pol
C 40	25.4	9.3	262	9	ACH44670 Human foe
C 41	25.4	9.3	264	2	AAZ07150 Human lun
C 42	25.2	9.2	109	12	ACH86303 Human gen
C 43	25.2	9.2	138	2	AAx11982 Human bia
C 44	25.2	9.2	219	8	ABx76158 Lung canc
C 45	25.2	9.2	219	11	ADP64955 Human sma

ALIGNMENTS

RESULT 1

ABQ55864/c

ID ABQ55864 standard; cDNA; 192 BP.

AC ABQ55864;

DT 22-AUG-2002 (first entry)

XX Human ovarian antigen HPAMQ76 cDNA, SEQ ID NO:1744.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;  
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
XX PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
XX inflammatory condition; immune disorder; blood disorder;  
XX cardiovascular disorder; respiratory disorder; neurological disorder;  
XX gastrointestinal disorder; urinary system disorder; drug screening;  
XX gene therapy; chromosome mapping; forensic analysis;  
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
XX antiinflammatory; gynaecological; reproductive; gene; ss.

OS Homo sapiens.

XX WO200200677-A1.

PD 03-JAN-2002.

PF 07-JUN-2001; 2001WO-US018569.

PR 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

DR P-PSDB; ABP42787.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
PT cancer), immune disorders, cardiovascular disorders and neurological  
XX diseases.

PS Claim 1; SEQ ID NO 1744; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigen (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired  
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
CC respiratory disorders, neurological disorders, gastrointestinal disorders  
CC and urinary system disorders. Ovarian antigen polypeptides and  
CC polynucleotides may also be used in screening for compounds which  
CC modulate ovarian antigen expression or activity. The polynucleotides may  
CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents cDNA encoding a human ovarian antigen of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 192 BP; 41 A; 57 C; 58 G;36 T; 0 U; 0 Other;

SQ	Sequence	192 BP; 41 A; 57 C; 58 G; 36 T; 0 U; 0 Other;
	Query Match	10.9%; Score 29.8; DB 6; Length 192;
	Best Local Similarity	58.4%; Pred. No. 85;
	Matches	52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY	63	ACCTCTGTATACCTTTGAGATGGCGGTCCGAGATGGCAGCAAAATTCGCAACTGCTGGG 122
Db	128	ACTTTCTGTAGCTTCAGACGCACGACCTTAGGGCAGCAGGACCACCATCGCTTTT 69
QY	123	GTGGGCTCTGGGTCCGTTGGAGGGCGGCA 151
Db	68	TCCTGTGCTAGGCGGTGGGATGCCGTCA 40

RESULT 2  
ACH82956/c  
ID ACH82956 standard; DNA: 265 BP.

XX  
AC ACH82956;

29-JUL-2004 (first entry)

DE Human genome derived single exon probe #16151.

Human; probe; ss; gene expression; single exon probe; microarray;  
alternative splicing event; genomic alteration.

OS Homo sapiens.

XX  
PN  
US2003194704-A1.

16-OCT-2003.

03-APR-2002; 2002US-00029386.

03-APR-2002; 2002US-00029386.

PA (PENN/) PENN S G.

PA (RANK//) RANK D R.

PA (HANZ/) HANZEL D K.

XX

PI	Penn SG, Rank DR, Hanzel DK;
XX	
DR	WPI; 2004-119264/12.
XX	
PT	New human genome-derived single exon nuc
PT	gene expression analysis, for identifying
PT	splicing events, for assessing genomic a
PT	surveying tissues.
XX	
PS	Claim 1; SEQ ID NO 16151; 80pp; English.

Claim 1; SEQ ID NO 16151; 80pp; English.

The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, a method of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human, single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html?DocID=20030194704](http://seqdata.uspto.gov/sequence.html?DocID=20030194704)

SQ Sequence 265 BP; 51 A; 97 C; 60 G; 57 T; 0 U; 0 Other;

Query Match 10.5%; Score 28.6; DB 12; Length 265;

Best Local Similarity 49.0%; Pred. No. 2.1e+02;

2008 Local Summary 15.0%; RICU NO. 2.1E+02;  
 Matches 76; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 2 TGGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCCAGC 61

Db - 252 TGGAGATAGAGCTTCCACAGGGCTCTCTTCATGTCTCTGGTTCCTCAGGCTGTAGATGAGC 193

62 TACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCA

### RESULT 3

RESUL 3  
ABN20543

ID ABN20543 standard; cDNA; 194 BP.

AC ABN20543;

DT 24-JUN-2002 (first entry)



CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX  
SQ Sequence 122 BP; 17 A; 31 C; 39 G; 35 T; 0 U; 0 Other;  
Query Match 10.1%; Score 27.6; DB 12; Length 122;  
Best Local Similarity 58.5%; Pred. No. 3.3e+02;  
Matches 48; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
XX  
QY 128 CTCTGGTGGTGGAGGGCGGAGTCTCGGCATGTAGTCTTCAGGTTCTGCGAGGG 187  
Db 1 CTAGATTTCAGGTCAATTCGCGCATGATCGCTGTGCAAGCCAGGTTCTTACAGGA 60  
QY 188 CTGAGGAAAGGCTGTGAGTGG 209  
Db 61 ACGTACCGGCGTCTGCTG 82  
RESULT 5  
ADQ91957/c  
ID ADQ91957 standard; DNA; 133 BP.  
XX  
AC ADQ91957;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE Dog dopamine receptor D4, DRD4, intron 2 long form.  
XX  
KW Dog; DRD4; intron 2; dopamine receptor D4; dog candidate; seeing-eye dog;  
KW disaster rescue dog; assisting dog; narcotics detection dog; police dog;  
KW animal herding dog; hunting dog; watchdog; housedog; ds.  
XX  
OS Canis familiaris.  
XX  
FH Key Location/Qualifiers  
FT allele 50..67  
FT /\*tag= a  
FT /note= "This region is deleted in the short form"  
XX  
PN US2004121368-A1.  
XX  
PD 24-JUN-2004.  
XX  
XX 15-AUG-2003; 2003US-00641428.  
XX  
PR 24-DEC-2002; 2002JP-00373006.  
XX  
PA (UYGI-) UNIV GIFU.  
XX  
XX Murayama M, Ito S;  
XX  
XX WPI; 2004-479668/45.  
XX  
PT New polynucleotides and polypeptides, useful for screening useful dog  
PT candidates having genetic aptitudes beneficial to people, or for  
PT screening probes or identifying probes used in southern hybridization or  
PT in situ hybridization.  
XX  
PS Claim 14; SEQ ID NO 8; 13pp; English.  
XX  
CC The invention relates to a polynucleotide derived from dog dopamine  
CC receptor D4 (DRD4) comprising the short and long forms of exon 1  
CC (ADQ91952 and ADQ91954) or the deleted portion from the long form  
CC (ADQ91950), and the short and long forms of intron 2 (ADQ91957 and  
CC ADQ91958 or the deleted portion of the long form (ADQ91956). Also  
CC included are a transformant in which a polynucleotide is expressively  
CC introduced that contains a nucleotide sequence selected from ADQ91950,  
CC ADQ91952, ADQ91954, ADQ91956, ADQ91957 and ADQ91958, a polypeptide  
CC containing an amino acid sequence encoded by the above nucleic acids  
CC (ADQ91953 or ADQ91955, screening for useful dog candidates, a kit for  
CC screening for useful dog candidates comprising a pair of PCR primers that  
CC bonds to the vicinity of both terminals of an allele in exon 1 or intron

CC 2 of a dog dopamine receptor D4 gene, breeding useful dog candidates that  
CC uses a dog in which the genotype of the allele in exon 1 or intron 2 of a  
CC dog dopamine receptor D4 gene is homozygotic for at least one of the  
CC parent dogs and a useful dog candidate that is selected based on the  
CC genotype of the allele in exon 1 or intron 2 of a dog dopamine receptor  
CC D4 gene. The service dogs are seeing-eye dogs, disaster rescue dogs,  
CC assisting dogs, narcotics detection dogs, police dogs, dogs for the  
CC hearing impaired, animal herding dogs, or hunting dogs. The dogs with  
CC genetic aptitudes beneficial to people are watchdogs or housedogs. Dogs  
CC having the short form of exon 1 are suited to being seeing eye dogs, dogs  
CC having the short form of intron 2 display mild-mannered behavioural  
CC traits and dogs having the long form of intron 2 display aggressive  
CC behavioural traits. The polynucleotides and polypeptides are useful for  
CC screening useful dog candidates having genetic aptitudes beneficial to  
CC people. The polynucleotides can be used as screening probes or  
CC identifying probes used in southern hybridisation, northern  
CC hybridization, or in situ hybridisation. The present sequence is the Dog  
CC dopamine receptor D4, DRD4, intron 2 long form.  
XX  
SQ Sequence 133 BP; 7 A; 83 C; 35 G; 8 T; 0 U; 0 Other;  
Query Match 10.0%; Score 27.2; DB 12; Length 133;  
Best Local Similarity 53.8%; Pred. No. 4.4e+02;  
Matches 56; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
XX  
QY 137 GGTGGAGGGCGGCGAGTCTCGGCATGTAGTCTTCAGGTTCTGCGAGGGTCCGAGGA 196  
Db 116 GGGCGGCGGGCGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 57  
QY 197 AGCGTGTCTACGTCGCGCTGAGATTGTCAGCGCGGGTCCCGAGGC 240  
Db 56 GGGCGGTCGCGAGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 13  
RESULT 6  
ADF31067/c  
ID ADF31067 standard; DNA; 235 BP.  
XX  
AC ADF31067;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Soil metagenome glycosyl hydrolase homologue DNA SEQ ID NO 287.  
XX  
KW glycosyl hydrolase; amylase; enzyme; washing; cleaning; textile;  
KW raw material; cotton; starch liquefaction; ethanol production;  
KW oligosaccharide preparation; cyclodextrin; hydrolysis; food; animal feed;  
KW antistaling; additive; bread; adhesive; soil metagenome; ds.  
XX  
OS Synecocystis sp.  
XX  
PN WO2003054177-A2.  
XX  
XX 03-JUL-2003.  
XX  
PF 13-DEC-2002; 2002WO-EP014210.  
XX  
PR 21-DEC-2001; 2001DE-01063748.  
XX  
XX (HENK ) HENKEL KGAA.  
XX  
XX Breves R, Maurer K, Eck J, Lorenz P, Zinke H;  
XX  
XX WPI; 2003-559145/52.  
XX  
PT New glycosyl hydrolase, useful for degrading starch, for e.g. in washing  
PT compositions, and in preparation of foods and animal feeds, e.g. as  
PT antistaling additives for bread.  
XX  
PS Claim 34; SEQ ID NO 287; 275pp; German.  
XX  
CC This invention describes novel glycosyl hydrolases and their derivatives  
CC which have amylase activity and are used (i) for fusing proteins



CC least an immunogenic part of a tumour protein are disclosed. The  
CC polypeptides are useful for inhibiting the development of cancer,

SQ sequence 264 BP; 52 A; 69 C; 63 G; 74 T; 0 U; 6 Other;





```
RESULT 13
AAI54432
ID AAI54432 standard; DNA; 167 BP.
XX AC
XX AAI54432;
XX DT
XX 17-OCT-2001 (first entry)
XX DE
XX Probe #23118 used to measure gene expression in human placenta sample.
XX KW
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX KW
XX Genetic disorder; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200157272-A2.
XX PD
XX 09-AUG-2001.
XX PF
XX 30-JAN-2001; 2001WO-US000663.
XX PR
XX 04-FEB-2000; 2000US-0180312P.
XX PR
XX 26-MAY-2000; 2000US-0207456P.
XX PR
XX 30-JUN-2000; 2000US-00608408.
XX PR
XX 03-AUG-2000; 2000US-00632366.
XX PR
XX 21-SEP-2000; 2000US-0234687P.
XX PR
XX 27-SEP-2000; 2000US-0236359P.
XX PR
XX 04-OCT-2000; 2000GB-00024263.
XX PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX PD
XX WPI; 2001-488897/53.
XX PS
XX Claim 25; SEQ ID NO 23118; 654pp; English.
XX CC
XX The present invention relates to single exon nucleic acid probes useful for analyzing
XX CC
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX CC
XX gene expression in human placenta.
XX PS
XX Claim 25; SEQ ID NO 23118; 654pp; English.
XX CC
XX The present invention relates to single exon nucleic acid probes (SNP).
XX CC
XX The present sequence is one such probe. The probes are useful for
XX CC
XX producing a microarray for predicting, measuring and displaying gene
XX CC
XX expression in samples derived from human placenta. The probes are useful
XX CC
XX for antenatal diagnosis of human genetic disorders
XX PS
XX Sequence 167 BP; 19 A; 53 C; 52 G; 43 T; 0 U; 0 Other;
XX SQ
Query Match 9.7%; Score 26.4; DB 4; Length 167;
Best Local Similarity 65.0%; Pred. No. 8e+02;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 114 CCTGCTGGGCTTGGCTCGGTCGTTGGAGGGCGGACGTCTCGGCATGTAGTCTTC 173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 CCTGCTGGGCTTGGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 84
RESULT 14
ABA39062
ID ABA39062 standard; DNA; 167 BP.
XX AC
XX ABA39062;
XX DT
XX 23-JAN-2002 (first entry)
XX DE
XX Probe #17528 for gene expression analysis in human heart cell sample.
XX KW
XX Human; gene expression; heart; microarray; vascular system; probe;
XX KW
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW
XX congenital heart disease; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200157276-A2.
XX PD
XX 09-AUG-2001.
XX PF
XX 30-JAN-2001; 2001WO-US000668.
XX PR
XX 04-FEB-2000; 2000US-0180312P.
XX PR
XX 26-MAY-2000; 2000US-0207456P.
XX PR
XX 30-JUN-2000; 2000US-00608408.
XX PR
XX 03-AUG-2000; 2000US-00632366.
XX PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX PD
XX WPI; 2001-488897/53.
XX PS
XX Claim 25; SEQ ID NO 23118; 654pp; English.
XX CC
XX The present invention relates to single exon nucleic acid probes (SNP).
XX CC
XX The present sequence is one such probe. The probes are useful for
XX CC
XX producing a microarray for predicting, measuring and displaying gene
XX CC
XX expression in samples derived from human placenta. The probes are useful
XX CC
XX for antenatal diagnosis of human genetic disorders
XX PS
XX Sequence 167 BP; 19 A; 53 C; 52 G; 43 T; 0 U; 0 Other;
XX SQ
Query Match 9.7%; Score 26.4; DB 4; Length 167;
Best Local Similarity 65.0%; Pred. No. 8e+02;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 114 CCTGCTGGGCTTGGCTCGGTCGTTGGAGGGCGGACGTCTCGGCATGTAGTCTTC 173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 CCTGCTGGGCTTGGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 84
RESULT 15
AAK48603
ID AAK48603 standard; DNA; 167 BP.
XX AC
XX AAK48603;
XX DT
XX 06-NOV-2001 (first entry)
XX DE
XX Human bone marrow expressed single exon probe SEQ ID NO: 23160.
XX KW
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW
XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200157276-A2.
XX PD
XX 09-AUG-2001.
XX PF
XX 30-JAN-2001; 2001WO-US000668.
XX PR
XX 04-FEB-2000; 2000US-0180312P.
XX PR
XX 26-MAY-2000; 2000US-0207456P.
XX PR
XX 30-JUN-2000; 2000US-00608408.
XX PR
XX 03-AUG-2000; 2000US-00632366.
XX PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX PD
XX WPI; 2001-488899/53.
XX PS
XX Claim 4; SEQ ID NO 17528; 530pp; English.
XX CC
XX The present invention relates to single exon nucleic acid probes for
XX CC
XX measuring human gene expression in a sample derived from human heart. The
XX CC
XX present sequence is one such probe. The probes may be used for
XX CC
XX predicting, measuring and displaying gene expression in samples derived
XX CC
XX from the human heart via microarrays. By measuring gene expression, the
XX CC
XX probes are useful for predicting, diagnosing, grading, staging,
XX CC
XX monitoring and prognosing diseases of the human heart and vascular system
XX CC
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC
XX congenital heart disease. Note: The sequence data for this patent did not
XX CC
XX form part of the printed specification, but was obtained in electronic
XX CC
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX PS
XX Sequence 167 BP; 19 A; 53 C; 52 G; 43 T; 0 U; 0 Other;
XX SQ
Query Match 9.7%; Score 26.4; DB 4; Length 167;
Best Local Similarity 65.0%; Pred. No. 8e+02;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 114 CCTGCTGGGCTTGGCTCGGTCGTTGGAGGGCGGACGTCTCGGCATGTAGTCTTC 173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 CCTGCTGGGCTTGGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 84
```

```

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GE-00024263.
XX
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 23160; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention
XX
XX SQ Sequence 167 BP; 19 A; 53 C; 52 G; 43 T; 0 U; 0 Other;

Query Match      9.7%; Score 26.4; DB 4; Length 167;
Best Local Similarity 65.0%; Pred. No. 8e+02;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 114 CCTGCTGGGCTTGGCTTGGCTTGGAGGGGCGGAGTCTGGCATGTAGTTC 173
    |||||
Db 25 CCTGCTGGGCTGGGCTGGGCTGGGCTGTCATTGTGCTGGCTGCTCTCTC 84

```

Search completed: January 31, 2005, 19:14:11  
Job time : 320 secs

***This Page Blank (uspto)***



```

RESULT 2
US-10-719-993-27866
; Sequence 27866.. Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27866
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-27866

```

	Query Match Best Local Similarity Matches	10.5%; 57.6%; 49;	Score 28.6; Pred. No. 27; Conservative 1;	Mismatches 35;	Indels 0;	Gaps 0;
QY	185	GGGCTGCAGAAAGGTGT	CAGTCGCCCTCAGATTGT	CAACGCGCGGGTCCCAGCCCTGC	244	
Dδ	70	GGGCAGCAGGAACA	CACACTGGCTTGGAAAGCAGGARGAT	TCCCAGAGGTGAGCTCTTGCCCTGC	129	
QY	245	ACCAGCTCACCAAGCTACGTTTCTCT	269			
Dδ	130	CTCTCCATGACTAGCTCTGTGCCT	154			

RESULT 3  
 US-10-029-386-16151/c  
 ; Sequence 16151, Application US/10029386  
 ; Publication No. US20030194704A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rankel, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
 ; FILE REFERENCE: AEOMICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029,386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288  
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 16151  
 ; LENGTH: 265  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO CHR19.1  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.4  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.8  
 ; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 6.3  
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.3  
 ; OTHER INFORMATION: NT HIT: G111276076, EVALUE 6.00e-05  
 ; OTHER INFORMATION: EST HUMAN HIT: BF062364.1, EVALUE 1.00e-03  
 ; OTHER INFORMATION: SWISSPROT HIT: Q9UGF6, EVALUE 5.00e-20  
 US-10-029-386-16151

	Query Match	10.5%	Score 28.6;	DB 15;	Length 265;
	Best Local Similarity	49.0%;	Pred. No. 28;		
	Matches	76;	Conservative	0;	Mismatches 79; Indels 0; Gaps 0;
QY	2	GGAGACACTACCGAAAGCTGGCTCTGTATAGATCCACGGCTTCCCAATGCCACG	61		
DB	252	GGAGATAGAGCTTCCACAGGGTCTCTTCATGCTGTGTCCTCAGGCTTAGATGAGC	193		

Qy	62	TACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGCGACGAAATTCGCAACCTGCTGG	121
Db	192	GGATTACGGGTGGGTATGAGGATGGAGTAGAAGACAGCAGCCAAAGCGCCCTTCACGGG	133
Qy	122	GGTTGGCTCTGGGTCGGTTCGATTGGAGGCGCGCAGTGT	156
Db	132	GAGTAGCGGGTGGTGGGTCGCATGTAGGTGGAGCT	98

```

RESULT 4
US-10-741-601-25556
; Sequence 25556, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25556
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-25556

```

	Query Match	10.4%	Score 28.4;	DB 17;	Length 201;
	Best Local Similarity	51.7%;	Prod. No. 31;		
	Matches 62;	Conservative 1;	Mismatches 57;	Indels 0;	Gaps 0
QY	97	GGCAGCAAAATTCGAAACCTCTGGGGTTGGCTCTCGGTCGGTTGGAGGGCGCAGTGCT	156		
Db	20	GGCGACAACAACCTCACCGAGATCGTGGGAGGCGAGAAATGCAAGACGGGAGTGTGCC	79		
QY	157	CGCGATGTAGTGTCTTCAGGTTCTGGCGAGGGCTGCAGGAAAGGCTCTCAGCTGCCTCTGAG	216		
Db	80	TGGCAGGTAAACAGTAGATATGTCCTCGGCGCTGCTGGAGAGACCACTTGGCCCGCTTG	139		

```

RESULT 5
US-10-242-535A-5913/c
; Sequence 5913, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5913
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-5913

```

```
Query Match      10.4%; Score 28.4; DB 16; Length 262;
Best Local Similarity 58.1%; Pred. No. 32;
Matches 50; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
```



QY 63 ACCTCTGATACCTTTCAGATGGGGTCCGAGATGGCAGCAAAATTTCGCAACTGCTGGG 122  
|||  
Db 86 ACTTCTTGTAGGCTTCAGACGACGACCTTGAGGGCAGCAGGAACACCATCGCTTTT 27  
|||  
QY 123 GTTGGCTCTGGTTCGGTTCGAGGCG 148  
|||  
Db 26 TCTGTCTAGGCGGTGGGATGCCG 1  
|||

## RESULT 6

US-10-085-783A-5913/c  
; Sequence 5913, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5913  
; LENGTH: 262  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-5913

Query Match 10.4%; Score 28.4; DB 16; Length 262;  
Best Local Similarity 58.1%; Pred. No. 32; Mismatches 0; Indels 0; Gaps 0;  
Matches 50; Conservative 0;

QY 63 ACCTCTGATACCTTTCAGATGGGGTCCGAGATGGCAGCAAAATTTCGCAACTGCTGGG 122  
|||  
Db 86 ACTTCTTGTAGGCTTCAGACGACGACCTTGAGGGCAGCAGGAACACCATCGCTTTT 27  
|||  
QY 123 GTTGGCTCTGGTTCGGTTCGAGGCG 148  
|||  
Db 26 TCTGTCTAGGCGGTGGGATGCCG 1  
|||

## RESULT 7

US-10-425-115-43846/c  
; Sequence 43846, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 43846  
; LENGTH: 177  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_139993C.1  
US-10-425-115-43846

Query Match 10.3%; Score 28; DB 18; Length 177;  
Best Local Similarity 51.6%; Pred. No. 42;  
Matches 64; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 38 CRAGCGCTTCCCAATGCCCCAGCTACCTCTGATACCTTTCAGATGCGGGTCCGAGATG 97  
|||  
Db 164 CACAGCAGCGCCATCGCCGCTGCAGACGCGCGCGGGGTAAATCGGTCCGCGATG 105  
|||  
QY 98 GCAGCAAAATTTCGAACCTGCTGGGTTGGCTCTGGGTGGGTTGGAGGGCGGAGTCTC 157  
|||  
Db 104 ACAACGACTACGAGGGCAATCGGGTTGGCTTTGGTTCGGATCGGTCTCTTAAGAGTACTC 45  
|||  
QY 158 GGCA 161  
|||  
Db 44 AGCA 41  
|||

## RESULT 8

US-10-767-701-21806/c  
; Sequence 21806, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 21806  
; LENGTH: 227  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(227)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 13390168  
US-10-767-701-21806

Query Match 10.3%; Score 28; DB 17; Length 227;  
Best Local Similarity 53.2%; Pred. No. 43;  
Matches 58; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 130 CTGGGTCTGGTTCGAGGCGCGAGTCTCGGCATGTAGTCTCAGGTTCTGCAGGGCT 189  
|||  
Db 212 CAGGGCCGTTGGGGTTCGCCGNGAGCGCCCTCGTGGGTCTCGGTGTCTCAGTCCGCCG 153  
|||  
QY 190 GCAGGAAAGCTGTACGTGCGCTGAGATTGTCAAGCGGCGGTCCCG 238  
|||  
Db 152 GCGGCTTAGGTTGGGGCGGCACGGTGCCTGGCAGGTTCGAGGAGCCAG 104  
|||

## RESULT 9

US-10-242-535A-24157/c  
; Sequence 24157, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994

SOFTWARE: PatentIn version 3.2  
SEQ ID NO 24157  
LENGTH: 154  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (122)..(122)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (122)..(122)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (135)..(135)  
OTHER INFORMATION: n is a, c, g, or t  
US-10-242-535A-24157

Query Match 10.2%; Score 27.8; DB 16; Length 154;  
Best Local Similarity 56.2%; Pred. No. 48;  
Matches 50; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
QY 63 ACTCTGTATACCCCTTGAGATGCGGGTCCGAGATGCGCAAAATTCGCAACCTGCTGGG 122  
DB 153 ACTTCTGTAGGCTTCANACGCCCGACCTTNAGGCGCAGCAGCCACCATCCGCTTTT 94  
QY 123 GTTGGCTCTGGTCCGTTGAGAGCGGCA 151  
DB 93 TCTTGTCTAGGCGGTGGATGCCGTCA 65

RESULT 10  
US-10-085-783A-24157/c  
Sequence 24157, Application US/10085783A  
Publication No. US20040037841A1  
GENERAL INFORMATION:  
APPLICANT: Liew, C.C.  
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
FILE REFERENCE: 4231/2002  
CURRENT APPLICATION NUMBER: US/10/085,783A  
CURRENT FILING DATE: 2002-02-28  
PRIOR APPLICATION NUMBER: US 60/305,340  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/275,017  
PRIOR FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: US 60/271,955  
PRIOR FILING DATE: 2001-02-28  
NUMBER OF SEQ ID NOS: 58994  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 24157  
LENGTH: 154  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (2)..(3)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (122)..(122)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (135)..(135)  
OTHER INFORMATION: n is a, c, g, or t  
US-10-085-783A-24157

Query Match 10.2%; Score 27.8; DB 16; Length 154;  
Best Local Similarity 56.2%; Pred. No. 48;  
Matches 50; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 63 ACTCTGTATACCCCTTGAGATGCGGGTCCGAGATGCGCAAAATTCGCAACCTGCTGGG 122  
DB 153 ACTTCTGTAGGCTTCANACGCCCGACCTTNAGGCGCAGCAGCCACCATCCGCTTTT 94  
QY 123 GTTGGCTCTGGTCCGTTGAGAGCGGCA 151  
DB 93 TCTTGTCTAGGCGGTGGATGCCGTCA 65

RESULT 11  
US-10-029-386-18496  
Sequence 18496, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: A60MICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 18496  
LENGTH: 122  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC000358.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8  
OTHER INFORMATION: SWISSPROT HIT: P46729, EVALUE 7.90e+00  
OTHER INFORMATION: EST HUMAN HIT: BE615665.1, EVALUE 1.50e-01  
OTHER INFORMATION: NT HIT: AF144523.1, EVALUE 1.00e-01  
US-10-029-386-18496

Query Match 10.1%; Score 27.6; DB 15; Length 122;  
Best Local Similarity 58.5%; Pred. No. 54;  
Matches 48; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
QY 128 CTCTGGTTCGGTTGGAGGCGGCGATGCTCGGCATGTAGTGTCTCAGGTTCTGCAGGG 187  
DB 1 CTAGGATTCAGTTCGAATTCGCGATGATCCCTGCTGCAAGCCAGGTTCTTACAGGA 60  
QY 188 CTGCGAGAAAGGCTGTCTAGCTG 209  
DB 61 ACGCTGACCGGGCTGTCTGCTG 82

RESULT 12  
US-10-741-601-19299  
Sequence 19299, Application US/10741601  
Publication No. US20040166519A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001500  
CURRENT APPLICATION NUMBER: US/10/741,601  
CURRENT FILING DATE: 2003-12-22  
NUMBER OF SEQ ID NOS: 26415  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19299  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-741-601-19299

Query Match 10.1%; Score 27.6; DB 17; Length 201;  
Best Local Similarity 60.8%; Pred. No. 57;  
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 198 GGCTGTGAGTGGCTGAGATTGTCAAGCGGGGTCCAGGGCTGCACAGGCTCACCAA 257  
| | | | |  
Db 110 GGCTGAGGTGATCGCTGCGCCACCATGCGGTGCTCCAGGACTGCACCCCATCTTCAA 169  
| | | | |  
QY 258 GCTACGTTTCCTTC 271  
| | | | |  
Db 170 GGTACTTATTAATC 183  
| | | | |

RESULT 13  
US-10-741-601-23785/c  
; Sequence 23785, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23785  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-23785

Query Match 10.1%; Score 27.6; DB 17; Length 201;  
Best Local Similarity 60.8%; Pred. No. 57; Mismatches 29; Indels 0; Gaps 0;  
Matches 45; Conservative 0;

QY 198 GGCTGTGAGTGGCTGAGATTGTCAAGCGGGGTCCAGGGCTGCACAGGCTCACCAA 257  
| | | | |  
Db 92 GGCTGAGGTGATCGCTGCGCCACCATGCGGTGCTCCAGGACTGCACCCCATCTTCAA 33  
| | | | |

QY 258 GCTACGTTTCCTTC 271  
| | | | |  
Db 32 GGTACTTATTAATC 19  
| | | | |

RESULT 14  
US-10-641-428-8/c  
; Sequence 8, Application US/10641428  
; Publication No. US20040121368A1  
; GENERAL INFORMATION:  
; APPLICANT: Gifu University  
; TITLE OF INVENTION: Polynucleotides, polypeptides and method for  
; TITLE OF INVENTION: screening for useful dog candidates.  
; FILE REFERENCE: P1P2002381US  
; CURRENT APPLICATION NUMBER: US/10/641,428  
; CURRENT FILING DATE: 2003-08-15  
; PRIOR APPLICATION NUMBER: JP2002-373006  
; PRIOR FILING DATE: 2002-12-24  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 133  
; TYPE: DNA  
; ORGANISM: Canis familiaris  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1)..(133)  
; OTHER INFORMATION: Dopamine Receptor D4 Intron 2 (Long form)  
US-10-641-428-8

Query Match 10.0%; Score 27.2; DB 17; Length 133;  
Best Local Similarity 53.8%; Pred. No. 74; Mismatches 48; Indels 0; Gaps 0;  
Matches 56; Conservative 0;

QY 137 GGTGAGGCGGCGAGTGTGCGCATGTAGTGTCTCAGGTTCTGGCAGGGCTGCAGGAA 196  
| | | | |  
Db 116 GGGCGGGCGGCGCGGGGCGCTGGCGAGGGGCGAGGGCGGGGCGCTGGCGAG 57  
| | | | |

QY 197 AGGCTGTGAGTGGCTGAGATTGTCAAGCGGGGTCCAGGCTGCACAGGCTCACCAA 240  
| | | | |  
Db 56 GGGCGGGTGGCGAGGGGGGCGCTGGCGGGGGGGGGGGCGGCGCAGCG 13  
| | | | |

RESULT 15  
US-10-872-874-287/c  
; Sequence 287, Application US/10872874  
; Publication No. US20050003419A1  
; GENERAL INFORMATION:  
; APPLICANT: Breves, Roland  
; APPLICANT: Maurer, Karl-Heinz  
; APPLICANT: Eck, Jurgen  
; APPLICANT: Lorenz, Patrick  
; APPLICANT: Zinke, Holger  
; TITLE OF INVENTION: GLYCOSYL HYDROLASES  
; FILE REFERENCE: HENK-0088 / H 5206  
; CURRENT APPLICATION NUMBER: US/10/872,874  
; CURRENT FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: PCT/EP02/14210  
; PRIOR FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: DE 101 63 748.9  
; PRIOR FILING DATE: 2001-12-21  
; NUMBER OF SEQ ID NOS: 320  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 287  
; LENGTH: 235  
; TYPE: DNA  
; ORGANISM: Synechocystis sp  
US-10-872-874-287

Query Match 10.0%; Score 27.2; DB 18; Length 235;  
Best Local Similarity 50.0%; Pred. No. 79; Mismatches 68; Indels 0; Gaps 0;  
Matches 68; Conservative 0;

QY 115 CTGCTGGGGTTGGCTCTGGGTGCGTTGGAGGGCGGCGAGTGTGCGCATGTAGTGTCTCA 174  
| | | | |  
Db 205 CTCCTTGGCCAGGTGACCGAGTGCAGCAGAGCTTGGCCATCATGCGGTTCTCGGT 146  
| | | | |

QY 175 GGTTCGGCAGGGCTGCAGGAAAGGCTGTGAGTGTGCGCTGAGATTGTCAAGCGGGGCTC 234  
| | | | |  
Db 145 GCGGTGTTTCGAGCAGCAGGTGCTGTTGGCCACCGAGCGGTGGCGGAGCAGCGCTGGTA 86  
| | | | |

QY 235 CCAGGCTGCACGAC 250  
| | | | |  
Db 85 GTAGCGGGCAGATC 70  
| | | | |

Search completed: January 31, 2005, 21:07:51  
Job time : 339 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 19:01:47 ; Search time 2458 Seconds  
(without alignments)  
4503.816 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_1\_273

Perfect score: 273

Sequence: 1 atggagcattaccggaaagc.....ccaagctactttcttcag 273

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 44308572 seqs, 20275418765 residues

Total number of hits satisfying chosen parameters: 55031126

Minimum DB seq length: 0

Maximum DB seq length: 273

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:\*

- 1: /cgn2\_6/ptodata/1/pna/PCTUS1 COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pna/PCTUS2 COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pna/US06 COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pna/US07 COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pna/US08 COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pna/US081 COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pna/US082 COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pna/US083 COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pna/US084 COMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pna/US085 COMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pna/US086 COMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pna/US087 COMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pna/US088 COMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pna/US089 COMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pna/US090 COMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pna/US091 COMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pna/US092A COMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pna/US092B COMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pna/US093A COMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pna/US093B COMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pna/US094 COMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pna/US095A COMB.seq:\*
- 23: /cgn2\_6/ptodata/1/pna/US095B COMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pna/US095C COMB.seq:\*
- 25: /cgn2\_6/ptodata/1/pna/US095D COMB.seq:\*
- 26: /cgn2\_6/ptodata/1/pna/US096A COMB.seq:\*
- 27: /cgn2\_6/ptodata/1/pna/US096B COMB.seq:\*
- 28: /cgn2\_6/ptodata/1/pna/US096C COMB.seq:\*
- 29: /cgn2\_6/ptodata/1/pna/US096D COMB.seq:\*
- 30: /cgn2\_6/ptodata/1/pna/US096E COMB.seq:\*
- 31: /cgn2\_6/ptodata/1/pna/US097A COMB.seq:\*
- 32: /cgn2\_6/ptodata/1/pna/US097B COMB.seq:\*
- 33: /cgn2\_6/ptodata/1/pna/US097C COMB.seq:\*
- 34: /cgn2\_6/ptodata/1/pna/US098A COMB.seq:\*
- 35: /cgn2\_6/ptodata/1/pna/US098B COMB.seq:\*
- 36: /cgn2\_6/ptodata/1/pna/US098C COMB.seq:\*
- 37: /cgn2\_6/ptodata/1/pna/US098D COMB.seq:\*
- 38: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*
- 39: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*
- 40: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*
- 41: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*
- 42: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*
- 43: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

- 44: /cgn2\_6/ptodata/1/pna/US099G COMB.seq:\*
- 45: /cgn2\_6/ptodata/1/pna/US100A COMB.seq:\*
- 46: /cgn2\_6/ptodata/1/pna/US100B COMB.seq:\*
- 47: /cgn2\_6/ptodata/1/pna/US101A COMB.seq:\*
- 48: /cgn2\_6/ptodata/1/pna/US101B COMB.seq:\*
- 49: /cgn2\_6/ptodata/1/pna/US102A COMB.seq:\*
- 50: /cgn2\_6/ptodata/1/pna/US102B COMB.seq:\*
- 51: /cgn2\_6/ptodata/1/pna/US103A COMB.seq:\*
- 52: /cgn2\_6/ptodata/1/pna/US103B COMB.seq:\*
- 53: /cgn2\_6/ptodata/1/pna/US104A COMB.seq:\*
- 54: /cgn2\_6/ptodata/1/pna/US104B COMB.seq:\*
- 55: /cgn2\_6/ptodata/1/pna/US105A COMB.seq:\*
- 56: /cgn2\_6/ptodata/1/pna/US105B COMB.seq:\*
- 57: /cgn2\_6/ptodata/1/pna/US106A COMB.seq:\*
- 58: /cgn2\_6/ptodata/1/pna/US107A COMB.seq:\*
- 59: /cgn2\_6/ptodata/1/pna/US107B COMB.seq:\*
- 60: /cgn2\_6/ptodata/1/pna/US107C COMB.seq:\*
- 61: /cgn2\_6/ptodata/1/pna/US107D COMB.seq:\*
- 62: /cgn2\_6/ptodata/1/pna/US108A COMB.seq:\*
- 63: /cgn2\_6/ptodata/1/pna/US108B COMB.seq:\*
- 64: /cgn2\_6/ptodata/1/pna/US109A COMB.seq:\*
- 65: /cgn2\_6/ptodata/1/pna/US109B COMB.seq:\*
- 66: /cgn2\_6/ptodata/1/pna/US109C COMB.seq:\*
- 67: /cgn2\_6/ptodata/1/pna/US109D COMB.seq:\*
- 68: /cgn2\_6/ptodata/1/pna/US109E COMB.seq:\*
- 69: /cgn2\_6/ptodata/1/pna/US109F COMB.seq:\*
- 70: /cgn2\_6/ptodata/1/pna/US109G COMB.seq:\*
- 71: /cgn2\_6/ptodata/1/pna/US109H COMB.seq:\*
- 72: /cgn2\_6/ptodata/1/pna/US109I COMB.seq:\*
- 73: /cgn2\_6/ptodata/1/pna/US109J COMB.seq:\*
- 74: /cgn2\_6/ptodata/1/pna/US109K COMB.seq:\*
- 75: /cgn2\_6/ptodata/1/pna/US109L COMB.seq:\*
- 76: /cgn2\_6/ptodata/1/pna/US109M COMB.seq:\*
- 77: /cgn2\_6/ptodata/1/pna/US109N COMB.seq:\*
- 78: /cgn2\_6/ptodata/1/pna/US109O COMB.seq:\*
- 79: /cgn2\_6/ptodata/1/pna/US109P COMB.seq:\*
- 80: /cgn2\_6/ptodata/1/pna/US109Q COMB.seq:\*
- 81: /cgn2\_6/ptodata/1/pna/US109R COMB.seq:\*
- 82: /cgn2\_6/ptodata/1/pna/US109S COMB.seq:\*
- 83: /cgn2\_6/ptodata/1/pna/US109T COMB.seq:\*
- 84: /cgn2\_6/ptodata/1/pna/US109U COMB.seq:\*
- 85: /cgn2\_6/ptodata/1/pna/US109V COMB.seq:\*
- 86: /cgn2\_6/ptodata/1/pna/US109W COMB.seq:\*
- 87: /cgn2\_6/ptodata/1/pna/US109X COMB.seq:\*
- 88: /cgn2\_6/ptodata/1/pna/US109Y COMB.seq:\*
- 89: /cgn2\_6/ptodata/1/pna/US109Z COMB.seq:\*
- 90: /cgn2\_6/ptodata/1/pna/US109AA COMB.seq:\*
- 91: /cgn2\_6/ptodata/1/pna/US109AB COMB.seq:\*
- 92: /cgn2\_6/ptodata/1/pna/US109AC COMB.seq:\*
- 93: /cgn2\_6/ptodata/1/pna/US109AD COMB.seq:\*
- 94: /cgn2\_6/ptodata/1/pna/US109AE COMB.seq:\*
- 95: /cgn2\_6/ptodata/1/pna/US109AF COMB.seq:\*
- 96: /cgn2\_6/ptodata/1/pna/US109AG COMB.seq:\*
- 97: /cgn2\_6/ptodata/1/pna/US109AH COMB.seq:\*
- 98: /cgn2\_6/ptodata/1/pna/US109AI COMB.seq:\*
- 99: /cgn2\_6/ptodata/1/pna/US109AJ COMB.seq:\*
- 100: /cgn2\_6/ptodata/1/pna/US109AK COMB.seq:\*
- 101: /cgn2\_6/ptodata/1/pna/US109AL COMB.seq:\*
- 102: /cgn2\_6/ptodata/1/pna/US109AM COMB.seq:\*
- 103: /cgn2\_6/ptodata/1/pna/US109AN COMB.seq:\*
- 104: /cgn2\_6/ptodata/1/pna/US109AO COMB.seq:\*
- 105: /cgn2\_6/ptodata/1/pna/US109AP COMB.seq:\*
- 106: /cgn2\_6/ptodata/1/pna/US109AQ COMB.seq:\*
- 107: /cgn2\_6/ptodata/1/pna/US109AR COMB.seq:\*
- 108: /cgn2\_6/ptodata/1/pna/US109AS COMB.seq:\*
- 109: /cgn2\_6/ptodata/1/pna/US109AT COMB.seq:\*
- 110: /cgn2\_6/ptodata/1/pna/US109AU COMB.seq:\*
- 111: /cgn2\_6/ptodata/1/pna/US109AV COMB.seq:\*
- 112: /cgn2\_6/ptodata/1/pna/US109AW COMB.seq:\*
- 113: /cgn2\_6/ptodata/1/pna/US109AX COMB.seq:\*
- 114: /cgn2\_6/ptodata/1/pna/US109AY COMB.seq:\*
- 115: /cgn2\_6/ptodata/1/pna/US109AZ COMB.seq:\*
- 116: /cgn2\_6/ptodata/1/pna/US109BA COMB.seq:\*

117: /cgn2\_6/ptodata/1/pna/us6051\_COMB.seq:  
118: /cgn2\_6/ptodata/1/pna/us6052\_COMB.seq:  
119: /cgn2\_6/ptodata/1/pna/us6053\_COMB.seq:  
120: /cgn2\_6/ptodata/1/pna/us6054\_COMB.seq:  
121: /cgn2\_6/ptodata/1/pna/us6055\_COMB.seq:  
122: /cgn2\_6/ptodata/1/pna/us6056\_COMB.seq:  
123: /cgn2\_6/ptodata/1/pna/us6057\_COMB.seq:  
124: /cgn2\_6/ptodata/1/pna/us6058\_COMB.seq:  
125: /cgn2\_6/ptodata/1/pna/us6059\_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	226	82.8	264	13	US-08-856-624-2456
2	226	82.8	264	24	US-09-540-766-8863
3	218	79.9	271	14	US-08-941-869A-2841
4	218	79.9	271	24	US-09-540-208-35269
5	212	77.7	245	9	US-08-438-571A-2389
6	212	77.7	245	24	US-09-540-233D-126181
7	206	75.5	260	14	US-08-959-395-768
8	206	75.5	260	23	US-09-539-334-6888
9	196	71.8	262	13	US-08-812-505-1415
10	196	71.8	262	14	US-08-978-620-1415
11	196	71.8	262	24	US-09-540-229-45100
12	196	71.8	262	66	US-60-013-529-1415
13	193	70.7	250	14	US-08-951-197-5067
14	193	70.7	250	23	US-09-539-806-30498
15	193	70.7	250	23	US-09-539-808B-30498
16	193	70.7	250	67	US-60-027-243-1037
17	192	70.3	250	23	US-09-539-800-16108
18	192	70.3	250	23	US-09-539-800B-16108
19	192	70.3	250	23	US-09-539-800C-16108
20	190	69.6	244	15	US-09-014-441-1305
21	190	69.6	244	24	US-09-540-212A-30093
22	190	69.6	244	58	US-60-034-841-1305
23	189.6	69.5	284	24	US-09-540-213-18789
24	185	67.8	239	14	US-08-901-902-73
25	185	67.8	239	23	US-09-539-806-30137
26	185	67.8	239	23	US-09-539-806B-30137
27	185	67.8	239	67	US-60-023-379-73
28	184	67.4	233	13	US-08-879-863-3710
29	184	67.4	233	13	US-08-879-863A-3710
30	184	67.4	233	24	US-09-540-213-12679
31	183.4	67.2	259	18	US-09-293-657-1609
32	183.4	67.2	259	24	US-09-540-233D-76641
33	183.4	67.2	259	73	US-60-082-035-1609
34	183	67.0	238	23	US-09-539-334-28833
35	181	66.3	238	24	US-09-540-229-113556
36	169	61.9	227	9	US-08-413-150-22
37	169	61.9	227	24	US-09-540-499-10618
38	169	61.9	227	24	US-09-540-499B-10618
39	166	60.8	224	23	US-09-539-800-15324
40	166	60.8	224	23	US-09-539-800B-15324
41	166	60.8	224	23	US-09-539-800C-15324
42	166	60.8	234	14	US-08-923-902-2187
43	166	60.8	234	24	US-09-540-208-7903
44	157	57.5	210	13	US-08-801-504-2052
45	157	57.5	210	24	US-09-540-499-16308

ALIGNMENTS

RESULT 1  
US-08-856-624-2456  
; Sequence 2456, Application US/08856624  
; GENERAL INFORMATION:

APPLICANT: Gooding, Douglas H.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Akerblom, Ingrid E.  
APPLICANT: Deleane, Angelo M.  
APPLICANT: Naughton, Rebecca E.  
APPLICANT: Klingler, Tod M.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
TITLE OF INVENTION: SIGMOID ADENOCARCINOMA  
NUMBER OF SEQUENCES: 3070  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/856,624  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/017,190  
FILING DATE: MAY 13, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,100  
FILING DATE: SEPTEMBER 25, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: CERRONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PD-0174 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 855-0555  
TELEFAX: (415) 845-4166  
INFORMATION FOR SEQ ID NO: 2456:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: 1929519  
US-08-856-624-2456

Query Match 82.8%; Score 226; DB 13; Length 264;  
Best Local Similarity 100.0%; Pred. No. 6.4e-51;  
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	48	CCCAATGCCAGCTACTCTCTGATACCTTGAGTGGGTCGAGATGCGACAAAT	107
Db	1	CCCAATGCCAGCTACTCTCTGATACCTTGAGTGGGTCGAGATGCGACAAAT	60
QY	108	TCGCAACCTGCTGGGGTTGGCTCTGGGTTCGAGGGCGGCGAGTGTCTCGCATGTAGT	167
Db	61	TCGCAACCTGCTGGGGTTGGCTCTGGGTTCGAGGGCGGCGAGTGTCTCGCATGTAGT	120
QY	168	GTTCCTCAGTCTTCGGCAGGGCTGCAGGAAGGCTGTGAGTGCCTGAGATTGTCAAGCG	227
Db	121	GTTCCTCAGTCTTCGGCAGGGCTGCAGGAAGGCTGTGAGTGCCTGAGATTGTCAAGCG	180
QY	228	CGCGGTCCAGGCTGCAGCAGCTCACCAGCTACGTTTCTTCAG	273
Db	181	CGCGGTCCAGGCTGCAGCAGCTCACCAGCTACGTTTCTTCAG	226

US-09-540-766-8863  
; Sequence 8863, Application US/09540766  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE  
; FILE REFERENCE: PD-1024 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,766  
; CURRENT FILING DATE: 2000-03-30  
; PRIOR APPLICATION DATA removed - refer to PALM or File Wrapper  
; NUMBER OF SEQ ID NOS: 77960  
; SOFTWARE: PERL Program  
; SEQ ID NO 8863  
; LENGTH: 264  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; FEATURE:  
; OTHER INFORMATION: Incyte ID No: hu00958497  
US-09-540-766-8863

Query Match 82.8%; Score 226; DB 24; Length 264;  
Best Local Similarity 100.0%; Pred. No. 6.4e-51;  
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 48 CCCAATGCCAGCTACCTCTGATACCTTGGATGCGGTCGAGATGCGCAGCAAAAT 107  
DB 1 CCCAATGCCAGCTACCTCTGATACCTTGGATGCGGTCGAGATGCGCAGCAAAAT 60  
  
QY 108 TCGCAACCTCTCGGGTGGCTCTGGTGGAGGGCGGAGTCTGGCATCTAGT 167  
DB 61 TCGCAACCTCTCGGGTGGCTCTGGTGGAGGGCGGAGTCTGGCATCTAGT 120  
  
QY 168 GTTCTCAGGTTCTGCGAGGGCTCAGGAAAGGCTGTGAGTGGCGCTGAGATTCTCAAGCG 227  
DB 121 GTTCTCAGGTTCTGCGAGGGCTCAGGAAAGGCTGTGAGTGGCGCTGAGATTCTCAAGCG 180  
  
QY 228 GCGGGTCCAGGCTGACACAGCTCACCAGCTACGTTTCCTTCAG 273  
DB 181 GCGGGTCCAGGCTGACACAGCTCACCAGCTACGTTTCCTTCAG 226

RESULT 3  
US-08-941-869A-2841  
; Sequence 2841, Application US/08941869A  
; GENERAL INFORMATION:  
; APPLICANT: Gooding, Douglas H.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Akerblom, Ingrid E.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Klingler, Tod M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
; TITLE OF INVENTION: HUMAN PROSTATE  
; NUMBER OF SEQUENCES: 5486  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

US-08-941-869A-2841  
; Sequence 2841, Application US/08941869A  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE  
; FILE REFERENCE: PD-1029 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,208  
; CURRENT FILING DATE: 2000-03-31  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 70811  
; SOFTWARE: PERL Program  
; SEQ ID NO 35269  
; LENGTH: 271  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-08-941-869A-2841

Query Match 79.9%; Score 218; DB 14; Length 271;  
Best Local Similarity 100.0%; Pred. No. 9.7e-49;  
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
DB 54 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 113  
  
QY 61 CTACCTCTCTGATACCTTGGATGCGGTCGAGATGCGGTCGAGATGCGCAAAATTCGCAACCTGCTG 120  
DB 114 CTACCTCTCTGATACCTTGGATGCGGTCGAGATGCGGTCGAGATGCGCAAAATTCGCAACCTGCTG 173  
  
QY 121 GGGTTGGCTCTCGGTCGGTTGGAGGGCGGAGTGTGAGTGGCGCTGCTGCGCATGTAGTGTTCAGGTTCT 180  
DB 174 GGGTTGGCTCTCGGTCGGTTGGAGGGCGGAGTGTGAGTGGCGCTGCTGCGCATGTAGTGTTCAGGTTCT 233  
  
QY 181 GGCAGGGCTGCAGGAAAGCTGTGAGTGGCGCTGAGAT 218  
DB 234 GGCAGGGCTGCAGGAAAGCTGTGAGTGGCGCTGAGAT 271

RESULT 4  
US-09-540-208-35269  
; Sequence 35269, Application US/09540208  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE  
; FILE REFERENCE: PD-1029 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,208  
; CURRENT FILING DATE: 2000-03-31  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 70811  
; SOFTWARE: PERL Program  
; SEQ ID NO 35269  
; LENGTH: 271  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

```
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00317250
US-09-540-208-35269

Query Match      79.9%; Score 218; DB 24; Length 271;
Best Local Similarity 100.0%; Pred. No. 9.7e-49;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGGCTTCCCAATGCCCCAG 60
Db 54 ATGAGCACATACCGGAAGCTGGCTCTGTAGAGCTCCAGGCTTCCCAATGCCCCAG 113
QY 61 CTACTCTCTGATACCTTGGAGATCGGGTCCGAGATCGGAGCAAAATTCGCAACCTGCTG 120
Db 114 CTACTCTCTGATACCTTGGAGATCGGGTCCGAGATCGGAGCAAAATTCGCAACCTGCTG 173
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 180
Db 174 GGGTTGGCTCTGGTTCGGTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 233
QY 181 GGCAGGCTGCAGGAAGGCTGTGCTGCTGCTGAGAT 218
Db 234 GGCAGGCTGCAGGAAGGCTGTGCTGCTGCTGAGAT 271

RESULT 5
US-08-438-571A-2389
; Sequence 2389, Application US/08438571A
; GENERAL INFORMATION:
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Levine, Wendy B.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Freitas, Nicole
; APPLICANT: Jernigan-Kelleher, Colleen
; APPLICANT: Stuart, Susan G.
; APPLICANT: Scott, Randal W.
; APPLICANT: Bills, Pamela Kay
; APPLICANT: Pham, Mino Thu
; APPLICANT: Altus, Christina M.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Akerman, Ingrid Erika
; TITLE OF INVENTION: POLYNUCLEOTIDES DERIVED FROM THP-1 CELLS
; NUMBER OF SEQUENCES: 5094
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,571A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C., Ph.D.
; REGISTRATION NUMBER: 39132
; REFERENCE/DOCKET NUMBER: PD-001-5 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 2389:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00448643
US-09-540-208-35269

Query Match      77.7%; Score 212; DB 9; Length 245;
Best Local Similarity 98.1%; Pred. No. 4.1e-47;
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGGCTTCCCAATGCCCCAG 60
Db 25 ATGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGGCTTCCCAATGCCCCAG 84
QY 61 CTACTCTCTGATACCTTGGAGATCGGGTCCGAGATCGGAGCAAAATTCGCAACCTGCTG 120
Db 85 CTACTCTCTGATACCTTGGAGATCGGGTCCGAGATCGGAGCAAAATTCGCAACCTGCTG 144
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 180
Db 145 GGGTTGGCTCTGGTTCGGTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 204
QY 181 GGCAGGCTGCAGGAAGGCTGTGCTGCTGCTGAGAT 216
Db 205 GGCAGGCTGCAGGAAGGCTGTGCTGCTGAGAT 240

RESULT 6
US-09-540-233D-126181
; Sequence 126181, Application US/09540233D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF HEMIC AND IMMUNE SYSTEM TISSUE
; FILE REFERENCE: PD-1030 CIP
; CURRENT APPLICATION NUMBER: US/09/540,233D
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/008,119
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 8/100,523
; PRIOR FILING DATE: 1993-08-03
; PRIOR APPLICATION NUMBER: 07/977,780
; PRIOR FILING DATE: 1992-11-19
; PRIOR APPLICATION NUMBER: 07/916,491
; PRIOR FILING DATE: 1992-07-17
; PRIOR APPLICATION NUMBER: 08/438,571
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/282,991
; PRIOR FILING DATE: 1994-07-28
; PRIOR APPLICATION NUMBER: 08/196,364
; PRIOR FILING DATE: 1994-02-14
; PRIOR APPLICATION NUMBER: 08/504,732
; PRIOR FILING DATE: 1995-07-20
; PRIOR APPLICATION NUMBER: 08/179,873
; PRIOR FILING DATE: 1994-01-11
; PRIOR APPLICATION NUMBER: 08/435,761
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 129340
; SOFTWARE: PERL Program
; SEQ ID NO 126181
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00448643
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 103, 105, 178, 231
```



OTHER INFORMATION: a, t, c, g, or other  
US-09-540-233D-126181

Query Match 77.7%; Score 212; DB 24; Length 245;  
Best Local Similarity 98.1%; Pred. No. 4.1e-47;  
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db |||||  
25 ATGAGCACTACCGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 84  
QY 61 CTACCTCTGATACCTTCCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db |||||  
85 CTACCTCTGATACCTTCCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 144  
QY 121 GGGTTGGCTCTGGGTCGGTGGAGGGCGGCGAGTGGCTCGGATGTAGTGTCTCAGGTTCT 180  
Db |||||  
145 GGGTTGGCTCTGGTGGTGGAGGGCGGCGAGTGGCTCGGATGTAGTGTCTCAGGTTCT 204  
QY 181 GGCAGGGCTCAGAAAGCTGTGAGTGGCTGAG 216  
Db |||||  
205 GGCAGGGCTCAGAAAGCTGTGAGTGGCTGAG 240

RESULT 7  
US-09-540-233D-126181  
; Sequence 768, Application US/08959395  
; GENERAL INFORMATION:  
; APPLICANT: Gooding, Douglas H.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Stuve, Susan G.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Akerblom, Ingrid E.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Klingler, Todd M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
; TITLE OF INVENTION: HUMAN PANCREATIC ISLET CELLS  
; NUMBER OF SEQUENCES: 5789  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/959,395  
; FILING DATE: HEREWITH  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/030,755  
; FILING DATE: OCTOBER 28, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/039,551  
; FILING DATE: DECEMBER 20, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CERRONE, MICHAEL C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PD-0259 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 768:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 260 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: 2071192H2  
US-08-959-395-768

Query Match 75.5%; Score 206; DB 14; Length 260;  
Best Local Similarity 100.0%; Pred. No. 1.8e-45;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db |||||  
49 ATGAGCACTACCGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 108  
QY 61 CTACCTCTGATACCTTCCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db |||||  
109 CTACCTCTGATACCTTCCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 168  
QY 121 GGGTTGGCTCTGGGTCGGTGGAGGGCGGCGAGTGGCTCGGATGTAGTGTCTCAGGTTCT 180  
Db |||||  
169 GGGTTGGCTCTGGTGGTGGAGGGCGGCGAGTGGCTCGGATGTAGTGTCTCAGGTTCT 228  
QY 181 GGCAGGGCTCAGAAAGCTGTCTCAG 206  
Db |||||  
229 GGCAGGGCTCAGAAAGCTGTCTCAG 254

RESULT 8  
US-09-539-334-6888  
; Sequence 6888, Application US/09539334  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Naughton, Sara J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF ENDOCRINE SYSTEM TISSUE  
; FILE REFERENCE: PD-1026 CIP  
; CURRENT APPLICATION NUMBER: US/09/539,334  
; CURRENT FILING DATE: 2000-03-30  
; "Prior application data removed - refer to PALM or file wrapper"  
; NUMBER OF SEQ ID NOS: 38381  
; SOFTWARE: PERL Program  
; SEQ ID NO 6888  
; LENGTH: 260  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: hu00887627  
US-09-539-334-6888

Query Match 75.5%; Score 206; DB 23; Length 260;  
Best Local Similarity 100.0%; Pred. No. 1.8e-45;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db |||||  
49 ATGAGCACTACCGAAAGCTGCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 108  
QY 61 CTACCTCTGATACCTTCCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db |||||  
109 CTACCTCTGATACCTTCCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 168  
QY 121 GGGTTGGCTCTGGGTCGGTGGAGGGCGGCGAGTGGCTCGGATGTAGTGTCTCAGGTTCT 180  
Db |||||  
169 GGGTTGGCTCTGGTGGTGGAGGGCGGCGAGTGGCTCGGATGTAGTGTCTCAGGTTCT 228  
QY 181 GGCAGGGCTCAGAAAGCTGTCTCAG 206  
Db |||||  
229 GGCAGGGCTCAGAAAGCTGTCTCAG 254

```
RESULT 9
US-08-812-505-1415
; Sequence 1415, Application US/08812505
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 3896
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,505
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/013,529
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0126P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 852-0195
; INFORMATION FOR SEQ ID NO: 1415:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 862319
US-08-812-505-1415

Query Match 71.8%; Score 196; DB 13; Length 262;
Best Local Similarity 97.5%; Pred. No. 9.5e-43;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAGCTGGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 60
Db 60 ATGAGCACTACCGGANAGCTGGCTCTNTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 119
QY 61 CTACCTCTGATACCTTGGATGCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 120 CTACCTCTGATACCTTNNATGCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 179
QY 121 GGGTTGGCTCTGGGTGGAGGCGGCGAGTGTCCGCATGTAGTGTCTCAGGTTCT 180
Db 180 GGGTTGGCTCTGGGTGGAGGCGGCGAGTGTCCGCATGTAGTGTCTCAGGTTCT 239
QY 181 GGCAGGGCTGCAGGAAGGCT 201
Db 240 GGCAGGGCTGCAGGAAGGCT 260
```

```
RESULT 10
US-08-978-620-1415
; Sequence 1415, Application US/08978620
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 7797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,620
; FILING DATE: HEREWITH
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,151
; FILING DATE: DECEMBER 6, 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/812,505
; FILING DATE: MARCH 7, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0126-1 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 852-0195
; INFORMATION FOR SEQ ID NO: 1415:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 862319
US-08-978-620-1415

Query Match 71.8%; Score 196; DB 14; Length 262;
Best Local Similarity 97.5%; Pred. No. 9.5e-43;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAGCTGGCTCTGTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 60
Db 60 ATGAGCACTACCGGANAGCTGGCTCTNTAGAGTCCCGAGCGCTTCCCAATGCCCGAG 119
QY 61 CTACCTCTGATACCTTGGATGCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 120 CTACCTCTGATACCTTNNATGCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 179
QY 121 GGGTTGGCTCTGGGTGGAGGCGGCGAGTGTCCGCATGTAGTGTCTCAGGTTCT 180
Db 180 GGGTTGGCTCTGGGTGGAGGCGGCGAGTGTCCGCATGTAGTGTCTCAGGTTCT 239
```

Qy 181 GGCAGGGCTGCAGGAAGGCT 201  
|||  
Db 240 GGCAGGGCTGCAGGAAGGCT 260

RESULT 11  
US-09-540-229-45100  
; Sequence 45100; Application US/09540229  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS  
; FILE REFERENCE: PD-1033 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,229  
; CURRENT FILING DATE: 2000-03-31  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 193582  
; SOFTWARE: PERL Program  
; SEQ ID NO 45100  
; LENGTH: 262  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: hu00691274  
; NAME/KEY: unsure  
; LOCATION: 38, 76, 87, 138, 140, 168  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-540-229-45100

RESULT 12  
US-60-013-529-1415  
; Sequence 1415, Application US/60013529  
; GENERAL INFORMATION:  
; APPLICANT: Gooding, Douglas H.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Akerblom, Ingrid E.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Klingler, Tod M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
; TITLE OF INVENTION: ASTROCYTOMA  
; NUMBER OF SEQUENCES: 3896  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/60/013,529  
FILING DATE:

RESULT 13  
 US-08-351-197-5067  
 ; Sequence 5067, Application US/08951197  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gooding, Douglas H.  
 ; APPLICANT: Stuv, Laura L.  
 ; APPLICANT: Stuart, Susan G.  
 ; APPLICANT: Ito, Laura Y.  
 ; APPLICANT: Akerblom, Ingrid E.  
 ; APPLICANT: Delegeane, Angelo M.  
 ; APPLICANT: Naughton, Rebecca E.  
 ; APPLICANT: Klingler, Tod M.  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM HUMAN  
 ; TITLE OF INVENTION: BREAST  
 ; NUMBER OF SEQUENCES: 5970  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA

COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/951,197  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,249  
FILING DATE: OCTOBER 1, 1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/826,438  
FILING DATE: MARCH 20, 1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/822,285  
FILING DATE: MARCH 20, 1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: CERRONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PD-0143-1 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 5067:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 250 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
IMMEDIATE SOURCE:  
CLONE: 1997229  
US-08-951-197-5067

Query Match 70.7%; Score 193; DB 14; Length 250;  
Best Local Similarity 100.0%; Pred. No. 6.2e-42;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGAAAGCTGCTCTGTAGAGTCCCGAGATGGCGGCTCCCAAAATTCGCAACCTGCTG 120  
Db 58 ATGGAGCACTACCGAAAGCTGCTCTGTAGAGTCCCGAGATGGCGGCTCCCAAAATTCGCAACCTGCTG 117  
QY 61 CTACCTCTGATACCTTGTAGATGCGGGTCCGAGATGGCGGCTCCCAAAATTCGCAACCTGCTG 120  
Db 118 CTACCTCTGATACCTTGTAGATGCGGGTCCGAGATGGCGGCTCCCAAAATTCGCAACCTGCTG 177  
QY 121 GGGTTGGCTTGGTGGTGGAGGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
Db 178 GGGTTGGCTTGGTGGTGGAGGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 237  
QY 181 GGCAGGGCTGCAG 193  
Db 238 GGCAGGGCTGCAG 250

RESULT 14  
US-09-539-806-30498  
Sequence 30498, Application US/09539806  
GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE

FILE REFERENCE: PD-1027 CIP  
CURRENT APPLICATION NUMBER: US/09/539,806  
CURRENT Filing Date: 2000-03-30  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 48372  
SOFTWARE: PERL Program  
SEQ ID NO 30498  
LENGTH: 250  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: hu00555791  
US-09-539-806-30498

Query Match 70.7%; Score 193; DB 23; Length 250;  
Best Local Similarity 100.0%; Pred. No. 6.2e-42;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGAAAGCTGCTCTGTAGAGTCCCGAGATGGCGGCTCCCAAAATTCGCAACCTGCTG 60  
Db 58 ATGGAGCACTACCGAAAGCTGCTCTGTAGAGTCCCGAGATGGCGGCTCCCAAAATTCGCAACCTGCTG 117  
QY 61 CTACCTCTGATACCTTGTAGATGCGGGTCCGAGATGGCGGCTCCCAAAATTCGCAACCTGCTG 120  
Db 118 CTACCTCTGATACCTTGTAGATGCGGGTCCGAGATGGCGGCTCCCAAAATTCGCAACCTGCTG 177  
QY 121 GGGTTGGCTTGGTGGTGGAGGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
Db 178 GGGTTGGCTTGGTGGTGGAGGGCGGCGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 237  
QY 181 GGCAGGGCTGCAG 193  
Db 238 GGCAGGGCTGCAG 250

RESULT 15  
US-09-539-806B-30498  
Sequence 30498, Application US/09539806B  
GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE  
FILE REFERENCE: PD-1027 CIP  
CURRENT APPLICATION NUMBER: US/09/539,806B  
CURRENT Filing Date: 2000-03-30  
PRIOR APPLICATION NUMBER: 08/706,766  
PRIOR Filing Date: September 27, 1996  
PRIOR APPLICATION NUMBER: 60/004,676  
PRIOR Filing Date: October 2, 1995  
PRIOR APPLICATION NUMBER: 08/749,515  
PRIOR Filing Date: November 15, 1996  
PRIOR APPLICATION NUMBER: 60/006,810  
PRIOR Filing Date: November 15, 1995  
PRIOR APPLICATION NUMBER: 08/822,285  
PRIOR Filing Date: March 20, 1997  
PRIOR APPLICATION NUMBER: 60/013,696  
PRIOR Filing Date: March 20, 1996  
PRIOR APPLICATION NUMBER: 08/951,197  
PRIOR Filing Date: October 1, 1997  
PRIOR APPLICATION NUMBER: 60/027,249  
PRIOR Filing Date: October 1, 1996  
PRIOR APPLICATION NUMBER: 08/826,438  
PRIOR Filing Date: March 20, 1997  
PRIOR APPLICATION NUMBER: 60/016,145  
PRIOR Filing Date: April 18, 1996  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 48372  
SOFTWARE: PERL Program

SEQ ID NO 30498  
LENGTH: 250  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: hu0055791  
US-09-539-806B-30498

Query Match 70.7%; Score 193; DB 23; Length 250;  
Best Local Similarity 100.0%; Pred. No. 6.2e-42;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGTCCCGCGCTTCCCAATGCCCCAG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
58 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGTCCCGCGCTTCCCAATGCCCCAG 117  
Qy 61 CTACCTCCTGATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
118 CTACCTCCTGATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 177  
Qy 121 GGGTTGGCTCTGGGTGCGTTGGAGGGGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
178 GGGTTGGCTCTGGGTGCGTTGGAGGGGGGCGGAGTGTCTGGCATGTAGTGTCTCAGGTTCT 237  
Qy 181 GGCAGGCTGCAG 193  
Db ||||||||||||  
238 GGCAGGCTGCAG 250

Search completed: January 31, 2005, 20:58:10  
Job time : 2461 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 19:09:11 ; Search time 74 Seconds  
(without alignments)

2622.235 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_1\_273

Perfect score: 273

Sequence: 1 atggagcataccggaagc.....ccaagctacgttcttcag 273

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1158410

Minimum DB seq length: 0

Maximum DB seq length: 273

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	9.9	264	US-09-370-838-7	Sequence 7, Appli
2	27	9.9	264	US-09-854-133-7	Sequence 7, Appli
3	26	9.5	92	US-09-511-625B-45	Sequence 45, Appl
4	26	9.5	198	US-09-079-812B-30	Sequence 30, Appl
5	25.8	9.5	165	US-08-905-223-154	Sequence 154, App
6	25.6	9.4	212	US-09-313-294A-2448	Sequence 2448, Ap
7	25.6	9.4	261	US-09-489-039A-1139	Sequence 1139, Ap
8	25.4	9.3	194	US-09-513-999C-22180	Sequence 22180, A
9	25.4	9.3	235	US-09-016-434-661	Sequence 661, App
10	25	9.2	211	US-09-513-999C-14087	Sequence 14087, A
11	25	9.2	255	US-09-513-999C-36531	Sequence 36531, A
12	24.6	9.0	215	US-09-513-999C-2518	Sequence 2518, Ap
13	24.6	9.0	272	US-09-016-434-855	Sequence 855, App
14	24.4	8.9	95	US-09-513-999C-15097	Sequence 15097, A
15	24.4	8.9	261	US-09-489-039A-2953	Sequence 2953, Ap
16	24.2	8.9	163	US-09-621-976-7986	Sequence 7986, Ap
17	24.2	8.9	180	US-09-513-999C-22891	Sequence 22891, A
18	24	8.8	140	US-09-513-999C-20498	Sequence 20498, A
19	24	8.8	181	US-09-513-999C-20847	Sequence 20847, A
20	23.8	8.7	137	US-09-513-999C-9178	Sequence 9178, Ap
21	23.8	8.7	231	US-09-252-991A-2750	Sequence 2750, Ap
22	23.6	8.6	139	US-09-513-999C-17624	Sequence 17624, A
23	23.6	8.6	183	US-09-513-999C-15131	Sequence 15131, A
24	23.6	8.6	186	US-09-188-330-95	Sequence 95, Appl
25	23.6	8.6	186	US-09-312-283C-95	Sequence 95, Appl
26	23.6	8.6	234	US-08-420-235B-6	Sequence 6, Appli
27	23.6	8.6	234	US-08-793-624-6	Sequence 6, Appli

c	28	23.6	8.6	234	5	PCT-US95-10194-6	Sequence 6, Appli
	29	23.4	8.6	167	4	US-09-621-976-19219	Sequence 19219, A
	30	23.2	8.5	160	4	US-09-513-999C-12509	Sequence 12509, A
c	31	23.2	8.5	169	1	US-08-519-777-29	Sequence 29, Appl
c	32	23.2	8.5	169	1	US-08-742-035-29	Sequence 29, Appl
c	33	23.2	8.5	169	1	US-08-777-019-29	Sequence 29, Appl
c	34	23.2	8.5	169	2	US-08-777-143-29	Sequence 29, Appl
c	35	23.2	8.5	169	3	US-08-775-414-29	Sequence 29, Appl
c	36	23.2	8.5	169	3	US-08-931-858E-29	Sequence 29, Appl
c	37	23.2	8.5	169	3	US-08-981-739-29	Sequence 29, Appl
c	38	23.2	8.5	169	4	US-09-128-026-29	Sequence 29, Appl
c	39	23.2	8.5	169	4	US-09-220-616-29	Sequence 29, Appl
c	40	23.2	8.5	169	4	US-09-220-527-29	Sequence 29, Appl
c	41	23.2	8.5	169	4	US-09-220-407-29	Sequence 29, Appl
c	42	23.2	8.5	169	4	US-09-219-952-29	Sequence 29, Appl
c	43	23.2	8.5	192	4	US-09-489-039A-7126	Sequence 7126, Ap
c	44	23.2	8.5	246	4	US-09-489-039A-7074	Sequence 7074, Ap
c	45	23.2	8.5	251	4	US-09-313-294A-3512	Sequence 3512, Ap

ALIGNMENTS

RESULT 1

US-09-370-838-7  
; Sequence 7, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; FILE REFERENCE: 210121.475C1  
; CURRENT APPLICATION NUMBER: US/09/370,838  
; EARLIER FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/285,323  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 264  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(264)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-370-838-7

Query Match 9.9%; Score 27; DB 4; Length 264;  
Best Local Similarity 54.3%; Pred. No. 42;  
Matches 51; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY	175	GGTTCGTGGCAGGCTGCAGAAAGGCTGTGACGTGCGCTGAGATTGTCAAGCGCGCGGTC	234
Db	12	GGATGGGACGGGACACAGATMGACTTAACTGCCCCACCGTTTCMAGGAAGGATT	71

QY	235	CCAGGGCTGCACCAAGCTCACCAGCTACGTTTCC	268
Db	72	ACAGGGCTGAGCCACTGCGCGCGCTCTTCTCC	105

RESULT 2

US-09-854-133-7  
; Sequence 7, Application US/09854133  
; Patent No. 6759508  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.







```
RESULT 9
US-09-016-434-661
; Sequence 661, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 661:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT12
; CLONE: 2626716
US-09-016-434-661

Query Match          9.3%; Score 25.4; DB 4; Length 235;
Best Local Similarity 74.4%; Pred. No. 1.3e+02;
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 160 CATGTAGTCTTCAGGTTCTGCAGGGCTGCAGGAAGGCTG 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 CATCCAGTCTCCCACTTTCTATATGGGGTGCAGGAAGGCTG 101

RESULT 10
US-09-513-999C-14087/c
; Sequence 14087, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 2000-02-24
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 50/122,487
; PRIOR FILING DATE: 1999-02-26

Query Match          9.2%; Score 25; DB 4; Length 211;
Best Local Similarity 61.5%; Pred. No. 1.6e+02;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 174 AGGTTCTGCAGGGCTGCAGGAAGGCTGCTCAGCTGCCTGAGATTGTCAGCGCGGGT 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 AGGCCCCGGCTGGGATGCAGGGAGGGCCCGAAGGGGCTGAGCTCACCGAGCTCGGGCC 31
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 234 CCCAG 238
    ||| |||
Db 30 CCGAG 26
    ||| |||

RESULT 11
US-09-513-999C-36531/c
; Sequence 36531, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 2000-02-24
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 36531
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 35
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 60
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 132
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 148
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 152
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 161
; OTHER INFORMATION: h=a or c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 183
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 204
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 855:
SEQUENCE CHARACTERISTICS:
LENGTH: 272 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PGANOT01
CLONE: 619240
US-09-016-434-855

Query Match          9.0%; Score 24.6; DB 4; Length 272;
Best Local Similarity 48.3%; Pred. No. 2.3e+02;
Matches 69; Conservative 0; Mismatches 74; Indels 0; Gaps 0

QY      115 CTGCTGGGGTTGGCTCTGGGTGCGTTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCA 174
Db      9   CTGGCCCTGGCTGGCCGTGGGCAGCATGTGCCTCTTCATCGCCGCTTTGCGGTGGGCTGG 68

QY      175 GGTTCCTGGCAGGCTCGCAGGAAGGCTGTCTCAGTCTGGCTAGATGTCTCAAGCGCGGGTC 234
Db      69 GGGCCCATCCCCCTGGCTCTCATGTCTCAGATCTTCCCTCTGCATGTCAAGGGCGTGGCG 128

QY      235 CCAGGCGCTGCACGAGCTCACCAA 257
Db      129 ATAGGCATCTGGCTCTCACCAA 151

RESULT 14
US-09-513-999C-15097/c
; Sequence 15097, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15097
; LENGTH: 95
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-15097

Query Match          8.9%; Score 24.4; DB 4; Length 95;
Best Local Similarity 56.1%; Pred. No. 1.9e+02;
Matches 46; Conservative 0; Mismatches 36; Indels 0; Gaps 0

QY      158 GCATGTAGTGTCTCAGGTTCTGGCAGGCTCGCAGAAAGGCTGTCTCAGCTGCGCTGAGA 217
Db      88 GTCAATTTCACATCTCTTTTGGTGGTAATGCGTCAGTTAAGGCAGGAACGGGCGCATCTGG 29

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 855:
SEQUENCE CHARACTERISTICS:
LENGTH: 272 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PGANOT01
CLONE: 619240
US-09-016-434-855

Query Match          9.0%; Score 24.6; DB 4; Length 272;
Best Local Similarity 48.3%; Pred. No. 2.3e+02;
Matches 69; Conservative 0; Mismatches 74; Indels 0; Gaps 0

QY      115 CTGCTGGGGTTGGCTCTGGGTGCGTTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCA 174
Db      9   CTGGCCCTGGCTGGCCGTGGGCAGCATGTGCCTCTTCATCGCCGCTTTGCGGTGGGCTGG 68

QY      175 GGTTCCTGGCAGGCTCGCAGGAAGGCTGTCTCAGTCTGGCTAGATGTCTCAAGCGCGGGTC 234
Db      69 GGGCCCATCCCCCTGGCTCTCTCATGTCTCAGATCTTCCCTCTGCATGTCTCAAGGGCGTGGCG 128

QY      235 CCAGGCGCTGCACGAGCTCACCAA 257
Db      129 ATAGGCATCTGGCTCTCACCAA 151

RESULT 14
US-09-513-999C-15097/c
Sequence 15097, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961
FILE REFERENCE: 59.US2.REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 15097
LENGTH: 95
TYPE: DNA
ORGANISM: Homo sapiens
US-09-513-999C-15097

Query Match          8.9%; Score 24.4; DB 4; Length 95;
Best Local Similarity 56.1%; Pred. No. 1.9e+02;
Matches 46; Conservative 0; Mismatches 36; Indels 0; Gaps 0

QY      158 GCATGTAGTGTCTCAGGTTCTGGCAGGCTCGCAGAAAGGCTGTCTCAGCTGCGCTGAGA 217
Db      88 GTCAATTTCACATCTCTTTTGGTGGTAATGCGTCAGTTAAGGCAGGAACGGGCGCATCTGCG 29

```



Result No.	Query			DB	ID	Description
	Score	Match	Length			
C 1	25.8	21.0	97	6	C0461900	Sequence
C 2	24.4	19.8	100	6	AX998871	Sequence
C 3	23.6	19.3	101	11	EV185660	EV185660 sqgm14870
C 4	23.4	19.0	89	4	AJ629187	AJ629187 Sus scrofa
C 5	23.4	19.0	100	9	S70266	S70266 TAP1 (TAP1-
C 6	23.4	19.0	100	9	S70274S1	S70274 TAP1 (TAP1-
C 7	23.4	19.0	101	9	HSITGAD02	U04275 Human beta-
C 8	23	18.7	89	6	AX899650	Sequence
C 9	23	18.7	89	6	BD035183	Sequence
C 10	23	18.7	121	6	AX262946	Sequence
C 11	23	18.7	121	6	AX262947	Sequence
C 12	22.6	18.4	90	6	A69898	AX69898 Sequence 27
C 13	22.6	18.4	90	6	BD022719	BD022719 Mirror pl
C 14	22.2	18.0	117	6	CQ665784	Sequence
C 15	22	17.9	106	3	PCREP2D	Z54253 P.clarkii m
C 16	21.8	17.7	75	3	AF400169	AF400169 Eisenia a
C 17	21.8	17.7	100	6	AX995957	Sequence
C 18	21.8	17.7	100	9	S70256S1	S70256 TAP1 (TAP1-
C 19	21.6	17.6	75	10	S60901	S60901 TCR V beta1

```
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
  1. .100
  /organism="Escherichia coli"
  /mol_type="unassigned DNA"
  /db_xref="taxon:562"
  /notes="mrec b3251 U00096 complement (3396512__3397615)"
ORIGIN
  Query Match 19.8%; Score 24.4; DB 6; Length 100;
  Best Local Similarity 58.1%; Pred. No. 1.8e+04;
  Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 5 AGCTAGCTTTCTTCCAGACTGAGGACAGCTGGGTCCAGCCTCACCTGCACAGGGCTAG 64
Db 27 AGCTACTGGCGCTGCTGGCGGACACAGCCTGCTGCTATTCGCTCTCCACAAAGGGCTAC 86
QY 65 ACCCCCTCAGTG 78
Db 87 GCCGCGCGCAAGTG 100
RESULT 3
LOCUS
DEFINITION
  sqm148706 Human DNA (Sequenc) Homo sapiens STS genomic, sequence
  tagged site.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
  Contact: Andreas Braun
  Pharmaceuticals division
  Sequenom, Inc.
  3595 John Hopkins Court, San Diego, CA 92121, USA
  Tel: 18582029018
  Fax: 18582029020
  Email: abraun@sequenom.com
  Primer A: No primer sequence submitted
  Primer B: No primer sequence submitted
  STS size: 101.
  Location/Qualifiers
  1. .101
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /clone_lib="Human DNA (Sequenc)"
  <1..>101
STS
ORIGIN
  Query Match 19.3%; Score 23.8; DB 11; Length 101;
  Best Local Similarity 58.0%; Pred. No. 2.8e+04;
  Matches 40; Conservative 1; Mismatches 28; Indels 0; Gaps 0;
QY 2 CCAAGCTAGTTTCTTCCAGACTGAGGACAGCTGGGTCCAGCCTCACCTGCACAGGGC 61
Db 29 CCAGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 88
QY 62 TAGACCCCCC 70
Db 89 CAGAGGTCC 97
RESULT 4
LOCUS
DEFINITION
  AJ629187
  Sus scrofa partial NFYA gene for nuclear transcription factor Y
  alpha.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
  AJ629187
  Sus scrofa partial NFYA gene for nuclear transcription factor Y
  alpha.
  AJ629187.1 GI:44903520
  NFYA gene; nuclear transcription factor Y alpha.
  Sus scrofa (pig)
  Sus scrofa
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
  Barbossa,A., Demeure,O., Urien,C., Milan,D., Chardon,P. and
  Renard,C.
  A physical map of large segments of pig chromosome 7q1.1-q1.4:
  comparative analysis with human chromosome 6p21
  Unpublished
  Barbossa,A.
  Direct Submission
  Submitted (27-FEB-2004) Barbossa A., Animal Genetic, Institut
  National de Recherche Agronom., LREG, INRA/CEA, domaine de Vilvert,
  78352 Jouy-en-Josas cedex, FRANCE
  Location/Qualifiers
  1. .89
  /organism="Sus scrofa"
  /mol_type="genomic DNA"
  /db_xref="taxon:9823"
  /chromosome="SSC7q11"
  /haplotype="H01"
  /clone_lib="SBAB genomic BAC library"
  74. .89
  /gene="NFYA"
  74..>89
  /gene="NFYA"
  /codon_start=1
  /product="nuclear transcription factor Y alpha"
  /protein_id="CAF32808.1"
  /db_xref="GI:44903521"
  /translation="MWSIQ"
ORIGIN
  Query Match 19.0%; Score 23.4; DB 4; Length 89;
  Best Local Similarity 60.0%; Pred. No. 3.7e+04;
  Matches 39; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 44 CTTCACTGACACAGGGCTAGACCCCTCAGCTGAGTGGCGGCATGTCCTGAGTGTGG 103
Db 76 CATTGCTGCCACTGCTAGTGTGTCACCGTCAGATTCCTTGCCCACTGCTGTTGTGG 17
QY 104 TCCTG 108
Db 16 TGTG 12
RESULT 5
LOCUS
DEFINITION
  S70266
  TAP1 (TAP1-B) {ATP binding region} [human, Caucosoids, mRNA
  Partial, 100 nt].
ACCESSION
VERSION
  S70266
  S70266.1 GI:546540
```

```
AX998871.1 GI:41005217
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1
Donner,H., Drescher,B., Huber,A. and Weber,J.
Biochip
Patent: EP 1260592-A 10334 27-NOV-2002;
MWG -Biotech AG (DE)
Location/Qualifiers
1. .100
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
/notes="mrec b3251 U00096 complement (3396512__3397615)"
ORIGIN
  Query Match 19.8%; Score 24.4; DB 6; Length 100;
  Best Local Similarity 58.1%; Pred. No. 1.8e+04;
  Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 5 AGCTAGCTTTCTTCCAGACTGAGGACAGCTGGGTCCAGCCTCACCTGCACAGGGCTAG 64
Db 27 AGCTACTGGCGCTGCTGGCGGACACAGCCTGCTGCTATTCGCTCTCCACAAAGGGCTAC 86
QY 65 ACCCCCTCAGTG 78
Db 87 GCCGCGCGCAAGTG 100
RESULT 3
LOCUS
DEFINITION
  sqm148706 Human DNA (Sequenc) Homo sapiens STS genomic, sequence
  tagged site.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
  Contact: Andreas Braun
  Pharmaceuticals division
  Sequenom, Inc.
  3595 John Hopkins Court, San Diego, CA 92121, USA
  Tel: 18582029018
  Fax: 18582029020
  Email: abraun@sequenom.com
  Primer A: No primer sequence submitted
  Primer B: No primer sequence submitted
  STS size: 101.
  Location/Qualifiers
  1. .101
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /clone_lib="Human DNA (Sequenc)"
  <1..>101
STS
ORIGIN
  Query Match 19.3%; Score 23.8; DB 11; Length 101;
  Best Local Similarity 58.0%; Pred. No. 2.8e+04;
  Matches 40; Conservative 1; Mismatches 28; Indels 0; Gaps 0;
```

```
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 100)
AUTHORS      Szafer, F., Oksenberg, J.R. and Steinman, L.
TITLE        New allelic polymorphisms in TAP genes
JOURNAL      Immunogenetics 39 (5), 374 (1994)
MEDLINE      94222452
PUBMED       8168860
REMARK       GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 148273] from the original journal article.
FEATURES
source      1..100
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
gene         1..100
             /gene="TAP1"
             /allele="TAP1-B"
CDS          1..100
             /gene="TAP1"
             /allele="TAP1-B"
             /note="Unknown"
             /codon_start=1
             /protein_id="AAD14055.1"
             /db_xref="GI:4261755"
             /translation="GYDTEVGEAGSQLSGGQQAVALARALKPCV"
ORIGIN
Query Match      19.0%; Score 23.4; DB 9; Length 100;
Best Local Similarity 73.2%; Pred. No. 3.6e+04;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 28 GACAGCTGGTCCAGCCTCACTGACACAGGCGCTAGACCC 68
    |||||
Db 41 GACAGCTGGTCCAGCCTCACTGACACAGGCGCTAGACCC 1

RESULT 6
S70274S1/c
LOCUS      TAP1 (TAP1-C) {ATP binding and 3' regions} (human, Caucasoids, mRNA
DEFINITION Partial, 100 nt, segment 1 of 2).
ACCESSION  S70274
VERSION     S70274.1 GI:546541
KEYWORDS
SEGMENT    1 of 2
SOURCE      Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 100)
AUTHORS      Szafer, F., Oksenberg, J.R. and Steinman, L.
TITLE        New allelic polymorphisms in TAP genes
JOURNAL      Immunogenetics 39 (5), 374 (1994)
MEDLINE      94222452
PUBMED       8168860
REMARK       GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 148274] from the original journal article.
FEATURES
source      1..100
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
gene         1..100
             /gene="TAP1"
             /allele="TAP1-C"
CDS          1..100
             /gene="TAP1"
             /allele="TAP1-C"
             /note="Unknown"

ORIGIN
Query Match      19.0%; Score 23.4; DB 9; Length 100;
Best Local Similarity 73.2%; Pred. No. 3.6e+04;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 28 GACAGCTGGTCCAGCCTCACTGACACAGGCGCTAGACCC 68
    |||||
Db 41 GACAGCTGGTCCAGCCTCACTGACACAGGCGCTAGACCC 1

RESULT 7
HSITGAD02/c
LOCUS      HSITGAD02
DEFINITION  Human beta-2 integrin alpha subunit (ITGAD) gene, partial exon 16,
and partial cds.
ACCESSION  U40275
VERSION     U40275.1 GI:1173860
KEYWORDS
SEGMENT     2 of 6
SOURCE      Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 101)
AUTHORS      Wong, D.A., Davis, E.M., LeBeau, M. and Springer, T.A.
TITLE        Cloning and chromosomal localization of a novel gene-encoding a
human beta 2-integrin alpha subunit
JOURNAL      Gene 171 (2), 291-294 (1996)
MEDLINE      96257236
PUBMED       8666289
REFERENCE    2 (bases 1 to 101)
AUTHORS      Wong, D.A.
TITLE        Direct Submission
JOURNAL      Submitted (07-NOV-1995) Dennis A. Wong, McMaster University,
Medicine, 1200 Main St. West, Hamilton, On L8N 3Z5, Canada
FEATURES
source      1..101
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="alpha W.1"
             /clone_lib="pWE15 Genomic"
             <1..4
             /gene="ITGAD"
             <5..>101
             /gene="ITGAD"
             /codon_start=1
             /product="beta-2 integrin alpha subunit"
             /protein_id="AAB60635.1"
             /db_xref="GI:1173867"
             /translation="RIASSQLSPRLQYFQALSGGDLTQDGLMDL"
             5..>101
             /gene="ITGAD"
             /number=16

ORIGIN
Query Match      19.0%; Score 23.4; DB 9; Length 101;
Best Local Similarity 60.0%; Pred. No. 3.6e+04;
Matches 39; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 41 CAGCCTCACTGACACAGGCTAGACCCCTCACAGTGGCGCCCATGTGCTGCAGTGT 100
    |||||
Db 91 CAGTCCATCTGGTGAGGTCTCTGACCCCTCCTGACCCCTCAGCGCTGTGCCAAATACATGCAGCCT 32

Qy 101 GGGTG 105
    |||
Db 31 GGGGG 27

ORIGIN
Query Match      19.0%; Score 23.4; DB 9; Length 101;
Best Local Similarity 60.0%; Pred. No. 3.6e+04;
Matches 39; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 41 CAGCCTCACTGACACAGGCTAGACCCCTCACAGTGGCGCCCATGTGCTGCAGTGT 100
    |||||
Db 91 CAGTCCATCTGGTGAGGTCTCTGACCCCTCCTGACCCCTCAGCGCTGTGCCAAATACATGCAGCCT 32

Qy 101 GGGTG 105
    |||
Db 31 GGGGG 27
```

```
RESULT 8
AX899650/c
LOCUS AX899650 89 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 15513 from Patent EP1033401.
ACCESSION AX899650
VERSION AX899650.1 GI:40054564
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.
AUTHORS Expressed sequence tags and encoded human proteins
TITLE Patent: EP 1033401-A 15513 06-SEP-2000;
JOURNAL Genset (IFR)
FEATURES
    source      Location/Qualifiers
    1..89
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
Query Match 18.7%; Score 23; DB 6; Length 89;
Best Local Similarity 60.3%; Pred. No. 4.8e+04;
Matches 38; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 60 GCTAGACCCCTCACAGTGGCGGCCATGTGCTGCAGTGTGGTGTCTCTCAGCCGGA 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 GCTCGTCCGCTCCGCGTGGCGGCGGCGGCTACTGCTGCTGCTCTCTCGGAT 11
QY 120 CCC 122
Db 10 GCC 8
RESULT 9
BD035183/c
LOCUS BD035183 89 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD035183
VERSION BD035183.1 GI:22576925
KEYWORDS JP 2001269182-A/11429.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 89)
Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
AUTHORS Sequence tag and encoded human protein
TITLE Patent: JP 2001269182-A 11429 02-OCT-2001;
JOURNAL GENSET
COMMENT OS Homo sapiens (human)
FN JP 2001269182-A/11429
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, BIMERIC DUCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09, C07K14/435, C07K16/19, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC
PH Key Location/Qualifiers.
FEATURES
    source      Location/Qualifiers
    1..89
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
ORIGIN
```

```
Query Match 18.7%; Score 23; DB 6; Length 89;
Best Local Similarity 60.3%; Pred. No. 4.8e+04;
Matches 38; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 60 GCTAGACCCCTCACAGTGGCGGCCATGTGCTGCAGTGTGGTGTCTCTCAGCCGGA 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 GCTCGTCCGCTCCGCGTGGCGGCGGCGGCTACTGCTGCTGCTCTCTCGGAT 11
QY 120 CCC 122
Db 10 GCC 8
RESULT 10
AX262946/c
LOCUS AX262946 121 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 337 from Patent WO0173002.
ACCESSION AX262946
VERSION AX262946.1 GI:16511745
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Kmiec, E.B., Gamper, H.B. and Rice, M.C.
AUTHORS Targeted chromosomal genomic alterations with modified single
TITLE stranded oligonucleotides
JOURNAL Patent: WO 0173002-A 337 04-OCT-2001;
UNIVERSITY OF DELAWARE (US)
FEATURES
    source      Location/Qualifiers
    1..121
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
Query Match 18.7%; Score 23; DB 6; Length 121;
Best Local Similarity 54.0%; Pred. No. 4.6e+04;
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 10 CGTTTCCTTACAGTCTGAGGACAGCTGGGTCCCGAGCTCACCTGACACAGGGCTAGACCCC 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 CGCTTAGTGTCTCTGGGGGAGCTCGTGGTGAGGCTCCCTTCTTTCGGGAGATCTCT 59
QY 70 CTCACAGTGGCGGCCCATGTGCTGCA 96
Db 58 TCCTCTGTGCGCGGCTCTCTCCAGGA 32
RESULT 11
AX262947
LOCUS AX262947 121 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 338 from Patent WO0173002.
ACCESSION AX262947
VERSION AX262947.1 GI:16511746
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Kmiec, E.B., Gamper, H.B. and Rice, M.C.
AUTHORS Targeted chromosomal genomic alterations with modified single
TITLE stranded oligonucleotides
JOURNAL Patent: WO 0173002-A 338 04-OCT-2001;
UNIVERSITY OF DELAWARE (US)
FEATURES
    source      Location/Qualifiers
    1..121
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
```



```
Query Match      18.7%; Score 23; DB 6; Length 121;
Best Local Similarity 54.0%; Pred. No. 4.6e+04;
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 10 CGTTTCTTCAGACTGAGGACGCTGGTCCAGCCTCACCTGACACAGGCTAGACCCC 69
    |||||
Db 4 CGTTTGTGTCCTCGGGGACGCTGCTGTGAGGCTCCCTTCTTGGGAGATTCTCT 63
    |||||
QY 70 CTCACAGTGGCGCCCATGTGCTGCA 96
    |||||
Db 64 TCTCTGTGCGCGGCTCTCCAGGA 90
    |||||

RESULT 12
A69898/c
LOCUS      90 bp      DNA      linear      PAT 07-MAY-1999
DEFINITION Sequence 27 from Patent WO9808856.
ACCESSION A69898
VERSION A69898.1 GI:4774410
KEYWORDS  .
SOURCE      unidentified
            unclassified.
ORGANISM    1 (bases 1 to 90)
REFERENCE   1 (bases 1 to 90)
AUTHORS    Bald,R., Erdmann,V.A. and Fuerste,J.P.
TITLE      MIRROR-SYMMETRICAL SELECTION AND EVOLUTION OF NUCLEIC ACIDS
JOURNAL    Patent: WO 9808856-A 27 05-MAR-1998;
            BALD ROLF (DE)
FEATURES    Location/Qualifiers
            source      1..90
                        /organism="unidentified"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:32644"
ORIGIN

Query Match      18.4%; Score 22.6; DB 6; Length 90;
Best Local Similarity 68.9%; Pred. No. 6.3e+04;
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2 CCAAGCTACGTTTCCTTCAGACTGAGGACGCTGGTCCAGCCT 46
    |||||
Db 90 CCAAGCTTGCATGCCTCGCAGACGCCGACAGCTCGTTCGCGGCT 46
    |||||

RESULT 14
CQ665784/c
LOCUS      117 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION Sequence 10710 from Patent WO02070737.
ACCESSION CQ665784
VERSION CQ665784.1 GI:42148013
KEYWORDS  .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS    Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE      Compositions and methods relating to osteoarthritis
JOURNAL    Patent: WO 02070737-A 10710 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES    Location/Qualifiers
            source      1..117
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
ORIGIN

Query Match      18.0%; Score 22.2; DB 6; Length 117;
Best Local Similarity 54.2%; Pred. No. 7.9e+04;
Matches 39; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 39 CCCAGCCTCACCTGCACACAGGCTAGACCCCTCACAGTCGCGCCGATGTGCTGCAGT 98
    |||||
Db 86 CCGTGCCCGTCTGCCGAGGCGAGGATCCCTCTCTGNTGCTCGTGCCNAATC 27
    |||||

QY 99 GTGGGTGCTGCT 110
    |||
Db 26 TTGGATCCNCT 15

RESULT 15
PCREP2D
LOCUS      106 bp      mRNA      linear      INV 14-JAN-2003
DEFINITION P.Clarkii mRNA; ac rich region (ID 2D).
ACCESSION 254253
VERSION 254253.1 GI:1001894
KEYWORDS  repeat.
SOURCE      Procamburus clarkii (red swamp crayfish)
ORGANISM    Procamburus clarkii
            Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
            Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
            Astacoidea; Cambaridae; Procamburus.
REFERENCE   1 (bases 1 to 106)
AUTHORS    Pekhleteki,R., Cooper,R.L., Atwood,H.L. and Hampson,D.R.
TITLE      Expression profiling of mRNA obtained from single identified
            crustacean motor neurons: determination of specificity of
            hybridization
JOURNAL    Invert. Neurosci. 1 (4), 341-349 (1996)
MEDLINE    98039421
PUBMED     9372152
REFERENCE   2 (bases 1 to 106)
            JOURNAL
            MEDLINE
            PUBMED
            REFERENCE
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 20:59:38 ; Search time 1963 Seconds  
(without alignments)  
2283.284 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_253\_375

Perfect score: 123

Sequence: 1 accaagctacgtttcttca.....tgctgtcagccggaccgcc 123

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 1226348

Minimum DB seq length: 0

Maximum DB seq length: 123

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	25.8	21.0	97	1	AA430669	AA430669 zw26d03.s
C 2	25.8	21.0	117	2	BF171548	BF171548 PCL2912.M
C 3	25.2	20.5	118	5	BY074656	BY074656 BY074656
C 4	24.8	20.2	120	1	AV631484	AV631484 AV631484
C 5	24.6	20.0	114	1	AA092420	AA092420 119365.se
C 6	24.4	19.8	112	8	AZ883677	AZ883677 RPCI-23-1
C 7	24.2	19.7	115	9	LBAR042F12	BX545265 Leishmani
C 8	24.2	19.5	52	1	AI262761	AI262761 Qk06e12.x
C 9	24.2	19.5	122	5	BQ972671	BQ972671 QHD4a07.y
C 10	23.8	19.3	90	1	AA079036	AA079036 zm94h12.i
C 11	23.8	19.3	110	7	N56039	N56039 J6239F.Huma
C 12	23.2	18.9	116	8	CC374455	CC374455 PUDM06TID
C 13	23.2	18.9	120	4	BJ504782	BJ504782 BJ504782
C 14	23.2	18.7	113	2	BE165373	BE165373 OV1-HT047
C 15	23.2	18.7	117	4	BJ472794	BJ472794 BJ472794
C 16	22.8	18.5	115	9	CC685407	CC685407 OGOAH39TV
C 17	22.8	18.5	117	2	BF291718	BF291718 WHE2205.B
C 18	22.6	18.4	115	9	CG709891	CG709891 1119015C1
C 19	22.6	18.4	122	9	CC886444	CC886444 SALK.1486
C 20	22.4	18.2	101	1	AI208489	AI208489 Q55504.x
C 21	22.4	18.2	104	8	AZ629277	AZ629277 IM0482D12
C 22	22.4	18.2	121	6	CB435325	CB435325 615414.MA
C 23	22.4	18.2	122	4	BM312764	BM312764 ig78g08.Y
C 24	22.4	18.2	123	2	BF956593	BF956593 PM1-NN120

25	22.4	18.2	123	2	BE171397	BE171397 RCI-HT054
26	22.2	18.0	79	6	CB298729	CB298729 220024.re
C 27	22.2	18.0	80	9	CL302865	CL302865 P015B08.G
C 28	22.2	18.0	88	9	CG869072	CG869072 AB0245.Sa
C 29	22.2	18.0	93	8	CC178569	CC178569 NPX449.Ba
C 30	22.2	18.0	102	4	BI046042	BI046042 MR3-FN020
C 31	22.2	18.0	109	1	AW000596	AW000596 614056B09
C 32	22.2	18.0	114	2	BE165383	BE165383 QV1-HT047
C 33	22.2	18.0	114	9	CG869066	CG869066 AB0239.Sa
C 34	22.2	18.0	119	6	BY641401	BY641401 BY641401
C 35	22.2	17.9	66	8	AZ656200	AZ656200 IM0531A10
C 36	22.2	17.9	110	4	BM484534	BM484534 538432.MA
C 37	22.2	17.9	111	4	BG274104	BG274104 WHE2231.F
C 38	22.2	17.9	120	1	AI617421	AI617421 zehnl542.
C 39	22.2	17.9	123	2	AW137016	AW137016 UI-H-B11-
C 40	22.2	17.7	123	8	BZ705049	BZ705049 PUCB34TD
C 41	21.8	17.7	91	1	AA639724	AA639724 nq93d05.9
C 42	21.8	17.7	113	9	CG477283	CG477283 OST1129.M
C 43	21.8	17.7	113	9	CG484766	CG484766 OST18836
C 44	21.8	17.7	116	2	BF019263	BF019263 UY03B04.Y
C 45	21.8	17.7	117	9	CC713422	CC713422 OGDH54TH

ALIGNMENTS

RESULT 1  
AA430669/c  
LOCUS AA430669 97 bp mRNA linear EST 16-OCT-1997  
DEFINITION zw26d03.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone  
IMAGE:770405 3', mRNA sequence.  
ACCESSION AA430669  
VERSION AA430669.1 GI:2111242  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 97)  
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubucque,T., Geisels,G., Jost,S.,  
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,  
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,  
White,Y., Wylie,T., Waterston,R. and Wilson,R.  
TITLE WashU-Merck EST Project 1997  
JOURNAL Unpublished (1997)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Possible reversed clone: polyt not found  
Seq primer: -41m13 fwd. ET from Amersham  
High quality sequence stop: 65.  
FEATURES  
Location/Qualifiers  
1..97  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5980279"  
/db\_xref="taxon:9606"  
/clone="IMAGE:770405"  
/sex="female"  
/tissue\_type="ovarian tumor"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares ovary tumor NbHOT"  
/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo (dT) primer [5'  
TGTTACCATCTGAAGTGGAGCGCGCGGTTTTTTTTTTT 3']  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

## ORIGIN

```

Query Match      21.0%; Score 25.8; DB 1; Length 97;
Best Local Similarity 56.5%; Pred. No. 4.7e+03;
Matches 48; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 39 CCAGGCTCACTGACACAGGCTAGACCCCTCAGAGTGGCGGCATGTGCTGCAGT 98
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 CCAGCGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 38
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 99 GTGGGTGCTGCTCAGCGCGGACCC 123
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37 ACCTCGCTGCTGCGGCGGACCC 13
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

## RESULT 2

```

BF171548/c
LOCUS
DEFINITION
PCL2912 Myeloma (PCL) cDNA library Homo sapiens cdna, mRNA
sequence.
BF171548
BF171548.1 GI:13437762
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 117)
Claudio J.O., Masih-Khan, E., Tang, H., Goncalves, J., Voralia, M.,
Li, Z.H., Nadeen, V., Cukerman, E., Francisco-Pabalan, O., Liew, C.C.,
Woodgett, J.R. and Stewart, A.K.
A molecular compendium of genes expressed in multiple myeloma
Blood 100 (6), 2175-2186 (2002)
22188429
12200383
Contact: A. Keith Stewart, M.D.
Oncology Research Network
University Health Network
610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
Tel: (416) 946-4639
Fax: (416) 946-6546
Email: k.stewart@utoronto.ca
PCR Primers
FORWARD: 5'-GCCAGCTGAATTAACTCACTCACTAAAGG-3'
BACKWARD: 5'-CCAGTGAATTGAATCACTCACTATAGGCG-3'
Seq primer: 5'-GAATTAACCTCACTAAAGG-3'

```

## FEATURES

```

1..117
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="male"
/tissue_type="Blood"
/cell_type="myeloma"
/dev_stage="Plasma cell leukemia"
/clone_lib="Myeloma (PCL) cDNA library"
/note="Vector: Lambda Zap Express; Site 1: EcoRI; Site 2:
XhoI; mRNA was purified from plasma cell leukemia
patient's peripheral blood containing >95% myeloma. An
oligo d(T)18 primer containing XhoI restriction site was
used to prime first strand synthesis using M-MLV reverse
transcriptase. To protect the cDNAs from XhoI digestion in
subsequent cloning step, the nucleotide analogue
5-methyl-dCTP was added to the nucleotide mixture and
[alpha-32p]dATP was added to monitor the quantity and quality
of first strand synthesis. After second-strand synthesis
and blunting of cDNA termini, EcoRI adapters were
ligated, followed by kinase treatment and digestion with
XhoI. The cDNAs were then size-fractionated using
Sephacryl S-500 column and then ligated into EcoRI and

```

XhoI digested Lambda Zap Express vector. The ligation product was packaged using Gigapack II packaging extract. The library had primary titre of approx. 1x10<sup>6</sup>. Clones from the primary library were randomly selected for single pass sequencing."

## ORIGIN

```

Query Match      21.0%; Score 25.8; DB 2; Length 117;
Best Local Similarity 56.5%; Pred. No. 4.8e+03;
Matches 48; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 11 GTTCTCTTCACTGAGGACAGCTGGGTCCAGCGCTCACTCAGAGGCTAGACCC 70
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101 GTCTGCTTAACACCAAGCTAGAGGGCGGGGCTCCCTCAAGCTGGGCTCGGCTCC 42
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 71 TCACAGTGGCGCGGCATGTGCTGC 95
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41 TCCCGCGCGGCGAGCTGGTGCCTCC 17
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

## RESULT 3

```

BY074656/c
LOCUS
DEFINITION
BY074656 RIKEN full-length enriched, 16 days embryo heart Mus
musculus cDNA clone I920096L06 5', mRNA sequence.
BY074656
BY074656.1 GI:26176407
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 118)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oeato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojibori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Flatcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Garcibol, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsukuba-shi, Ibaraki, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

```



47 TAGGTGTCTTTTGGSCGAGACAGGTGAGCCCTGGATCTCTTACACAGCGCAATAAG 106

58 CCC 70

107 ACC 109

RESULT 6	AZ883677/c	112 bp	DNA	linear	GSS 05-MAR-2001
LOCUS	RP01-23-187D12.TJ	RP01-23	Mus musculus	genomic clone	
DEFINITION	RP01-23-187D12, genomic survey sequence.				
ACCESSION	AZ883677				
VERSION	AZ883677.1	GI:13202622			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 112)				
AUTHORS	Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.				
TITLE	Muscle BAC End Sequences from Library RP01-23				
JOURNAL	Unpublished (1999)				
COMMENT	Other GSSs: RP01-23-187D12.TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208				

```

Query Match      19.88; Score 24.4; DB 8; Length 112;
Best Local Similarity 68.08; Pred. No. 1.2e+04;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

19 CAGACTGAGCAGCTGGGTCCGAGCCTCACCTGCACAGGGCTAGACCC 68
110 CAGACTGAGAAAAGATCTTTTCCAATCCTCATCAAAAAGGGCTCTATCC 61

```

RESULT 7			
LBAF042F12			
LOCUS	LBAF042F12	115 bp	DNA linear
			GSS 23-JUN-2003

```

DEFINITION Leishmania braziliensis GSS, clone LBAF042F12, genomic survey
sequence.
ACCESSION BX545265
VERSION BX545265.1 GI:32170063
KEYWORDS GSS; genomic survey sequence.
SOURCE Leishmania braziliensis
ORGANISM Leishmania braziliensis
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania; Leishmania braziliensis species complex.
REFERENCE 1
AUTHORS Laurentino,E.C., Ruiz,J.C. and Cruz,A.K.
TITLE GSS analysis of the Leishmania braziliensis genome
UNPUBLISHED
REFERENCE 2 (bases 1 to 115)
AUTHORS Cruz,A.K.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) Cruz A.K., University of Sao Paulo,
Department of Molecular and Cell Biology, FMRP, Avenida
Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL
COMMENT Clone requests: akcruz@fmrp.usp.br.
FEATURES             Location/Qualifiers
     source            1..115
                        /organism="Leishmania braziliensis"
                        /mol_type="genomic DNA"
                        /strain="MHOM/BR/75/M2904"
                        /db_xref="taxon:5660"
                        /clone="LBAF042F12"

```

## ORIGIN

Query Match	19.7%	Score 24.2;	DB 9;	Length 115;
Best Local Similarity	71.1%;	Pred. NO. 1.4e+04;		
Matches 32;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;
Qy	60	GCTAGACCCCTCACAGTGGCGGCCATGTGCTGCAGTGTGGGT	104	
Db	71	GCTAAGCCCTTCAGCTCGCAGACACCTGCGTGTAGTGTGGGT	115	

RESULT	8
AI262761	
LOCUS	AI262761 52 bp mRNA linear EST 13-NOV-1998
DEFINITION	qk06612.x1 NC1 CGAP Kid3 Homo sapiens cDNA clone IMAGE:1868206 3', similar to SW:PRP4.HUMAN P10163 SALIVARY PROLINE-RICH PROTEIN PO PRECURSOR ; contains PTR5..b3 MSR1 repetitive element ;, mRNA sequence.

Trace considered overall poor quality  
Seq primer: -40UP from Gibco  
High quality sequence stop: 1.  
Location/Qualifiers  
1. .52

## ORIGIN

19.58; Score 24; DB 5; Length 122;  
 All Similarity 60.9%; Pred. No. 1.6e+04;  
 39; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

28 GACAGCTGGGCTCCAGGCTCACCTGCACACAGGGCTAGACCCCTCTCAGTGCGCCCGCAT 87  
 85 GAATGTCATCTCCAGAGGGCGCTGACTCCAGGCTAGCCACCCAAACCGTGGTCCGGCAC 26

## 3

25 AGGC 22

Accession	Length	Source	Accession	Length	Source
AA079036	90 bp	mRNA	linear	EST 23-DEC-1997	
z99412.r1	Stratagene colon HT29 (#937221)				
IMAGE:545639	5', similar to gb:U7206				
UTMAM1	40S RIBOSOMAL PROTEIN S4				
	mRNA sequence				

(HUMAN) ; , mRNA sequence.  
AA078036

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 90)

Chissee, S., Dietrich, N., DuBouque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M. B., Tan, F., Thierry-Mieg, J., Treviskie, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R., and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)  
97044478  
8889549

000510  
Contact: Wilson RK  
Washington University School of Medicine

Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@atson.wustl.edu](mailto:est@atson.wustl.edu)  
WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.

```

this clone is available royalty-free through DMSU ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 1017 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 90
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3923828"
/db_xref="taxon:9606"
/clone="IWAGE:545639"
/dev_stage="HT-29 (HTB 38) adenocarcinoma"
/lab_host="SOuR (kanamycin resistant)"
/clone_lib="Stratagene colon HT29 (#937221)"
/note=vector: pBluescript SK-; Site 1: EcoRI Site 2:

```







***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 20:17:18 ; Search time 269 Seconds  
(without alignments)  
2400.293 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_253\_375

Perfect score: 123

Sequence: 1 accaagctacgtcttcttca.....tgctgctcagcgggacccc 123

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 4444388

Minimum DB seq length: 0  
Maximum DB seq length: 123

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_23Sep04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25.8	21.0	97	6	ABL81700 Human ova
C 2	25.2	20.5	121	10	ADP86506 Single nu
C 3	24.4	19.8	100	8	ACD79058 E. coli K
C 4	23.2	18.9	118	10	ADC40305 Human sec
C 5	23	18.7	89	3	AAC11438 Human sec
C 6	23	18.7	121	4	ABA77492
C 7	23	18.7	121	4	ABA77491
C 8	22.8	18.5	121	10	ADF87025
C 9	22.6	18.4	121	4	AAK70400
C 10	22.6	18.4	121	4	AAK70397
C 11	22.2	18.0	121	10	ADF87024
C 12	21.8	17.7	100	8	ACD76144
C 13	21.8	17.7	111	12	ACH89320
C 14	21.6	17.6	120	2	AAI12744
C 15	21.4	17.4	119	3	AAC11139
C 16	21.4	17.4	121	12	ADL80264
C 17	21.2	17.2	91	6	ABN63065
C 18	21.2	17.2	100	8	ACD79057
C 19	21.2	17.2	121	4	ABA77363
C 20	21.2	17.2	121	4	ABA77364
C 21	21	17.1	58	6	ABL32014

C 22	21	17.1	65	6	ABN54356
C 23	21	17.1	70	12	ADP05517
C 24	21	17.1	100	8	ACD70591
C 25	21	17.1	121	12	ADK91981
C 26	20.8	16.9	54	10	ADC35401
C 27	20.8	16.9	60	6	ABN59212
C 28	20.8	16.9	60	6	ABN59043
C 29	20.8	16.9	100	8	ACD72441
C 30	20.8	16.9	106	12	ACH85499
C 31	20.8	16.9	112	4	AAH34256
C 32	20.8	16.9	121	12	ADK91466
C 33	20.8	16.9	121	12	ADL80566
C 34	20.6	16.7	65	6	ABN57458
C 35	20.6	16.7	77	3	AAC20910
C 36	20.6	16.7	80	12	ADM96137
C 37	20.6	16.7	112	4	AAI24307
C 38	20.6	16.7	112	4	ABA69447
C 39	20.6	16.7	112	4	AAK43526
C 40	20.6	16.7	112	5	AAI09857
C 41	20.6	16.7	112	6	ABS17656
C 42	20.6	16.7	121	10	ADH92474
C 43	20.6	16.7	121	10	ADH92553
C 44	20.4	16.6	57	12	ADN35911
C 45	20.4	16.6	65	6	ABN53195

ALIGNMENTS

RESULT 1  
ABL81700/c  
ID ABL81700 standard; cDNA; 97 BP.  
XX  
AC ABL81700;  
XX  
DT 17-MAY-2002 (first entry)  
XX  
DE Human ovarian cancer related cDNA clone SEQ ID NO:4678.  
XX  
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200192581-A2.  
XX  
PD 06-DEC-2001.  
XX  
PF 29-MAY-2001; 2001WO-US017756.  
XX  
PR 26-MAY-2000; 2000US-0207484P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Algate PA, Harlocker SL, Jones R;  
XX  
DR WPI; 2002-122075/16.  
XX  
PT Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide.  
XX  
PS Claim 1; SEQ ID NO 4678; 489pp; English.  
XX  
CC The present invention describes a composition (I) comprising: carriers and immunostimulants; and a polypeptide (II) of a ovarian tumour CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1) from the 10912 nucleotide sequences as given in ABL77023 to ABL87934, CC (III) encoding (II) having a sequence (S2); a T cell population of (II), CC or antigen presenting cells that express (II). (I) has cytostatic CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for CC detecting ovarian cancer in a patient's biological sample preferably CC serum or ovarian tissue. The method comprises contacting a biological





PT encoding them useful for treating skin, neurodegenerative diseases, as an  
 PT antithrombotic agent and for inducing endothelial cell apoptosis.  
 XX  
 PS Example 41; SEQ ID NO 283; 473pp; English.

The invention discloses isolated PRO secreted/transmembrane polypeptides and the nucleic acid encoding them. The polypeptides can be used to raise antibodies that specifically bind to the PRO polypeptide, for linking a bioactive molecule to a cell expressing a PRO protein and for modulating at least one biological activity of a cell. PRO polypeptides are useful for detecting other PRO polypeptides in a sample and for linking a bioactive molecule to a cell expressing a PRO polypeptide. The PRO polypeptide antibodies are useful for modulating the biological activity of a cell expressing PRO polypeptides. The PRO polypeptides or polynucleotides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. These are useful for stimulating hypertrophy of neonatal heart, inhibiting vascular endothelial growth factor (VEGF)-stimulated proliferation of endothelial cells, modulating the proliferation of stimulated T-lymphocytes, enhancing the survival or proliferation of retinal neurons or rod photoreceptor cells, inducing c-fos in endothelial cells, modulating glucose or FFA uptake, inducing proliferation and/or re-differentiation of chondrocytes. In particular, these are useful for detecting or treating cardiac insufficiency disorders, wounds, cancerous tumours, retinal disorders or injuries (e.g. loss of sight due to retinitis pigmentosa), obesity, diabetes, hyperinsulinaemia, hypoinsulinaemia, or bone or cartilage disorders (e.g. sports injuries or arthritis) in mammals. PRO polypeptides and their portions affect the expression of genes which have a role in cell death. The polynucleotides are useful in molecular biology including uses as hybridisation probes for cDNA library to isolate the full-length PRO cDNA or to isolate other cDNAs, in chromosome and gene mapping, in the generation of antisense RNA and DNA, for preparing PRO polypeptides, for generating transgenic animals or knockout animals which are useful in the development and screening of therapeutically useful reagents, as probes and for the genetic analysis of individuals with genetic disorders as well as for recombinantly expressing the protein and for chromosome identification. The proteins are useful as molecular marker for protein electrophoresis purposes, as therapeutic agents, for screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). The polynucleotides and proteins are useful for tissue typing. PRO antibodies are useful for immunohistochemical staining and/or assay of sample fluids. Anti-PRO antibodies are useful in diagnostic assays for PRO e.g. detecting its expression in specific cells, tissues or serum and for affinity purification of PRO from recombinant cell culture or natural sources. The PRO genes may also be used in gene therapy, particularly for replacing a defective gene. The sequence presented is a PCR primer which was used to amplify a PRO polynucleotide of the invention.

SQ Sequence 118 BP; 25 A; 35 C; 35 G; 23 T; 0 U; 0 Other;  
Query Match 18.9%; Score 23.2; DB 10; Length 118;  
Best Local Similarity 53.3%; Pred. No. 1.3e+03;  
Matches 49; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

4 AACCTAGGTTTCCTTACAGCTGAGGACAGCTGGGTCCAGCCTCACCTGACACAGGGCTA 63  
23 AAGTTTCCTATCATCGACTCGGAAGTCTGAGCCATCTGTACTGGCGGGAGCAGGACAG 82  
64 GACCCCTCACAGTGGCGCGCCGATGTGCCTGC 95  
83 GGACCCATCACTCAGGACTGCTGTGTCCGCG 114

RESULT 5  
AAC11438/c  
ID AAC11438 standard; CDNA; 89 BP.  
XX  
XX  
XX AC  
XX AAC11438;  
XX  
XX  
XX  
XX 06-OCT-2000 (first entry)  
XX  
XX  
XX  
XX Human secreted protein 5' EST, SEQ ID NO: 15513.

XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW	gene therapy; chromosome mapping; ss.
XX	
OS	Homo sapiens.
XX	
FN	EP1033401-A2.
XX	
XX	06-SEP-2000.
PD	
XX	
XX	21-FEB-2000; 2000EP-00200610.
PP	
XX	
PR	26-FEB-1999; 99US-0122487P.
XX	
XX	(GEST ) GENSET.
PA	
XX	
PI	Dumas Milne Edwards J, Duclert A, Giordano J;
XX	
XX	WPI; 2000-500381/45.
DR	
XX	
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT	obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX	
PS	Claim 1; SEQ ID NO 15513; 7ipp + Sequence Listing; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors

SQ Sequence 89 BP; 19 A; 24 C; 40 G; 6 T; 0 U; 0 Other;  
 Query Match 18.7%; Score 23; DB 3; Length 89;  
 Best Local Similarity 60.3%; Pred. NO. 1.4e+03;  
 Matches 38; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 60 GCTAGACCCCTCACAGTGCGCCGCCATGTGCTCGAGTGTGGTGTCTCAGCCGGA 119  
 |||||  
 Db 70 GCTCGTCCCTCTCCGCGTGCACCGCCGCTACTGTGTGCTCTCTCCCGAT 11  
 |||||  
 QY 120 CCC 122  
 ||||  
 Db 10 GCC 8

RESULT 6	
ABA77492	
ID	ABA77492 standard; DNA; 121 BP.
XX	
XX	
XX	ABA77492;
XX	
DT	24-JAN-2002 (first entry)
XX	
DE	p53 mutation correcting oligonucleotide

Human, gene therapy; adenosine deaminase deficiency; p53; beta-globin; retinoblastoma; BRCA1; BRCA2; CFT8; cystic fibrosis; cancer; Factor V; cyclin-dependent kinase inhibitor 2A; CDKN2A; melanoma; APC; HBA1; HBA2; adenomatous polyposis of the colon; Factor VII; Factor IX; thrombosis; haemophilia; alpha thalassemia; haemoglobin alpha locus 1; MLH1; APOE; mismatch repair; MSH2; MSH6; hyperlipidaemia; apolipoprotein E; LDLR; familial hypercholesterolaemia; UCL1; syndrome; APP; PSEN1; antisenescence

KW UDP-glucuronosyltransferase; amyloid precursor protein; presenilin-1;  
 KW Alzheimer's disease; cytostatic; antisickling; antianaemic; haemostatic;  
 KW antileptic; ss.  
 XX  
 XX Homo sapiens.  
 XX WO200173002-A2.  
 XX 04-OCT-2001.  
 XX  
 XX 27-MAR-2001; 2001WO-US009761.  
 XX  
 XX 27-MAR-2000; 2000US-0192176P.  
 XX 27-MAR-2000; 2000US-0192176P.  
 XX 01-JUN-2000; 2000US-0208538P.  
 XX 30-OCT-2000; 2000US-0244989P.  
 XX  
 XX (UYDE ) UNIV DELAWARE.  
 XX  
 XX Kmiec EB, Gamper HB, Rice MC;  
 XX WPI; 2001-639230/73.  
 XX  
 XX Oligonucleotide for targeted alterations of genetic sequences and for  
 XX treating cystic fibrosis, comprises at least one mismatch and chemical  
 XX modification.  
 XX  
 XX Claim 7; Page 63; 294pp; English.  
 XX  
 XX The present invention provides single-stranded oligonucleotides which can  
 XX be used for the targeted alteration of genomic sequences, where the  
 XX oligonucleotide has at least one mismatch compared with the genomic  
 XX sequence to be altered. In particular, these sequences are directed at  
 XX the following genes: adenosine deaminase, p53, beta-globin,  
 XX retinoblastoma, BRCA1, BRCA2, CFTR, cyclin-dependent kinase inhibitor 2A  
 XX (CDKN2A), APC, Factor V, Factor VIII, Factor IX, haemoglobin alpha locus  
 XX 1 (HBA1), haemoglobin alpha locus 2 (HBA2), MLH1, MSH2, MSH6,  
 XX apolipoprotein E (APOE), LDL receptor (LDLR), UDP-glucuronosyltransferase  
 XX (UGT1), amyloid precursor protein (APC), presenilin-1 (PSEN1) and  
 XX presenilin-2 (PSEN2). These can be used in the gene therapy of diseases  
 XX such as cancer, adenosine deaminase deficiency, cystic fibrosis,  
 XX haemophilia, hypercholesterolaemia, thalassaemia, sickle cell anaemia,  
 XX Alzheimer's disease, melanoma, adenomatous polyposis of the colon and  
 XX various syndromes. The present sequence is one of the gene correcting  
 XX oligonucleotides of the invention  
 XX  
 XX Sequence 121 BP; 17 A; 43 C; 31 G; 30 T; 0 U; 0 Other;  
 XX  
 XX Query Match 18.7%; Score 23; DB 4; Length 121;  
 XX Best Local Similarity 54.0%; Pred. No. 1.5e+03;  
 XX Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
 QY 10 CGTTTCCTTCAGACTGAGGACAGCTGGTCCGAGCTCCTCAGCTACACAGGGCTAGACCCC 69  
 Db 4 CGCTTAGTGTCTCCCTGGGGGACGCTCGTGGTGAGGCTCCCTTTCTTGGCGGAGATTCTCT 63  
 QY 70 CTCACAGTGGCGCCGATGTCCTGCA 96  
 Db 64 TCCTCTGTGGCGGGTCTCTCCAGGA 90  
 RESULT 7  
 ABA77491/c  
 ID ABA77491 standard; DNA; 121 BP.  
 XX  
 XX ABA77491;  
 XX  
 XX 24-JAN-2002 (first entry)  
 XX  
 XX p53 mutation correcting oligonucleotide SEQ ID NO: 337.  
 XX  
 XX Human; gene therapy; adenosine deaminase deficiency; p53; beta-globin;  
 KW retinoblastoma; BRCA1; BRCA2; CFTR; cystic fibrosis; Cancer; Factor V;  
 KW cyclin-dependent kinase inhibitor 2A; CDKN2A; melanoma; APC; HBA1; HBA2;  
 KW adenomatous polyposis of the colon; Factor VII; Factor IX; thrombosis;  
 KW haemophilia; alpha thalassaemia; haemoglobin alpha locus 1; MLH1; APOE;  
 KW mismatch repair; MSH2; MSH6; hyperlipidaemia; apolipoprotein E; LDLR;  
 KW familial hypercholesterolaemia; UGT1; syndrome; APP; PSEN1; antisense;  
 KW UDP-glucuronosyltransferase; amyloid precursor protein; presenilin-1;  
 KW Alzheimer's disease; cytostatic; antisickling; antianaemic; haemostatic;  
 KW antileptic; ss.  
 XX  
 XX Homo sapiens.  
 XX WO200173002-A2.  
 XX 04-OCT-2001.  
 XX  
 XX 27-MAR-2001; 2001WO-US009761.  
 XX  
 XX 27-MAR-2000; 2000US-0192176P.  
 XX 27-MAR-2000; 2000US-0192176P.  
 XX 01-JUN-2000; 2000US-0208538P.  
 XX 30-OCT-2000; 2000US-0244989P.  
 XX  
 XX (UYDE ) UNIV DELAWARE.  
 XX  
 XX Kmiec EB, Gamper HB, Rice MC;  
 XX WPI; 2001-639230/73.  
 XX  
 XX Oligonucleotide for targeted alterations of genetic sequences and for  
 XX treating cystic fibrosis, comprises at least one mismatch and chemical  
 XX modification.  
 XX  
 XX Claim 7; Page 63; 294pp; English.  
 XX  
 XX The present invention provides single-stranded oligonucleotides which can  
 XX be used for the targeted alteration of genomic sequences, where the  
 XX oligonucleotide has at least one mismatch compared with the genomic  
 XX sequence to be altered. In particular, these sequences are directed at  
 XX the following genes: adenosine deaminase, p53, beta-globin,  
 XX retinoblastoma, BRCA1, BRCA2, CFTR, cyclin-dependent kinase inhibitor 2A  
 XX (CDKN2A), APC, Factor V, Factor VIII, Factor IX, haemoglobin alpha locus  
 XX 1 (HBA1), haemoglobin alpha locus 2 (HBA2), MLH1, MSH2, MSH6,  
 XX apolipoprotein E (APOE), LDL receptor (LDLR), UDP-glucuronosyltransferase  
 XX (UGT1), amyloid precursor protein (APC), presenilin-1 (PSEN1) and  
 XX presenilin-2 (PSEN2). These can be used in the gene therapy of diseases  
 XX such as cancer, adenosine deaminase deficiency, cystic fibrosis,  
 XX haemophilia, hypercholesterolaemia, thalassaemia, sickle cell anaemia,  
 XX Alzheimer's disease, melanoma, adenomatous polyposis of the colon and  
 XX various syndromes. The present sequence is one of the gene correcting  
 XX oligonucleotides of the invention  
 XX  
 XX Sequence 121 BP; 30 A; 31 C; 43 G; 17 T; 0 U; 0 Other;  
 XX  
 XX Query Match 18.7%; Score 23; DB 4; Length 121;  
 XX Best Local Similarity 54.0%; Pred. No. 1.5e+03;  
 XX Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
 QY 10 CGTTTCCTTCAGACTGAGGACAGCTGGTCCGAGCTCCTCAGCTACACAGGGCTAGACCCC 69  
 Db 118 CGCTTAGTGTCTCCCTGGGGGACGCTCGTGGTGAGGCTCCCTTTCTTGGCGGAGATTCTCT 59  
 QY 70 CTCACAGTGGCGCCGATGTCCTGCA 96  
 Db 58 TCCTCTGTGGCGGGTCTCTCCAGGA 32  
 RESULT 8  
 ADF87025/c  
 ID ADF87025 standard; DNA; 121 BP.  
 XX  
 XX ADF87025;  
 XX  
 XX 26-FEB-2004 (first entry)





PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249277P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483436/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX Disclosure; SEQ ID NO 25212; 3071pp + Sequence Listing; English.  
PS

XX  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 121 BP; 34 A; 19 C; 56 G; 12 T; 0 U; 0 Other;

Query Match 18.4%; Score 22.6; DB 4; Length 121;  
Best Local Similarity 60.7%; Pred. No. 2.1e+03;  
Matches 37; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 14 TCCTTCAGACTGAGGACAGCTGGTCCAGCCTCACCTGCACACAGGGCTAGACCCCTCA 73

Db 65 TGCCTCAGACTGACACTAGCTTCTTCACTGACCTCTCTGCTGAGGGCTGCCCTCTGA 6

QY 74 C 74  
|  
Db 5 C 5

## RESULT 10

AAK70397/c

ID AAK70397 standard; DNA; 121 BP.

XX AC AAK70397;

XX DT 06-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25209.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX OS Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001354.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.



CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 121 BP; 34 A; 19 C; 56 G; 12 T; 0 U; 0 Other;

Query Match 18.4%; Score 22.6; DB 4; Length 121;  
Best Local Similarity 60.7%; Pred. No. 2.1e+03;  
Matches 37; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 14 TCCTTCAGACTGAGCAGCTGGGTCCAGCCTCACCCTGACACAGGGCTAGACCCCTCA 73  
Db 65 TGCTTCAGACTGACACTAGTCTTCTCACTGACCTCTGCTGAGGGCTGCCCCCTGA 6

QY 74 C 74  
Db 5 C 5

## RESULT 11

ADF87024/c  
ID ADF87024 standard; DNA; 121 BP.

XX AC ADF87024;

XX DT 26-FEB-2004 (first entry)

XX Single nucleotide polymorphism detection human gene, SEQ ID No 607.

XX human; single nucleotide polymorphism; microarray; side effect; gene; ds.  
XX Homo sapiens.

XX JP2003235571-A.

XX PD 26-AUG-2003.

XX PF 12-FEB-2002; 2002JP-00034717.

XX PR 12-FEB-2002; 2002JP-00034717.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 2003-820454/77.

XX Novel polynucleotide useful for detecting single nucleotide polymorphisms in human gene.

XX Claim 1; SEQ ID NO 607; 704pp; Japanese.

XX The invention relates to a novel polynucleotide isolated and purified from a human gene having any one of 935 fully defined sequences as given in specification, or a sequence having a base substitution. The invention further relates to: an oligonucleotide containing single nucleotide polymorphisms; a PCR primer set chosen from the combination of two DNA fragments from any one of 1220 fully defined sequences as given in specification; a labelling probe containing the SNP containing oligo; and a microarray equipped with the SNP containing oligo. The isolated human gene of the invention is useful for detecting the single nucleotide polymorphisms in human gene. The isolated human gene is also useful for diagnosis of disease and determination of side effect to a medical agent. The isolated human gene is also effective in detecting single nucleotide polymorphisms in a human gene. This polynucleotide sequence represents one of the 935 isolated polynucleotides from a human gene of the invention.

SQ Sequence 121 BP; 24 A; 42 C; 40 G; 14 T; 0 U; 1 Other;

Query Match 18.0%; Score 22.2; DB 10; Length 121;

Best Local Similarity 77.1%; Pred. No. 2.8e+03;

Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 65 ACCCCCTCAGTGGCGCCCGCCATGTGCTGCATGTG 99

Db 57 ACCGCCCTCACATTGCTCCACCAGGTGCTGTGGGG 23

## RESULT 12

ACD76144

XX ID ACD76144 standard; DNA; 100 BP.

XX AC ACD76144;

XX DT 18-SEP-2003 (first entry)

XX E. coli K12 MG1655 biochip probe SEQ ID 7420.

XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.

XX Escherichia coli.

XX EP1260592-A1.

XX PD 27-NOV-2002.

XX PF 17-MAY-2001; 2001EP-00112179.

XX PR 17-MAY-2001; 2001EP-00112179.

XX (MWGB-) MWG-BIOTECH AG.

XX Donner H, Drescher B, Huber A, Weber J;

XX WPI; 2003-241155/24.

XX Biochip containing probes complementary with open reading frames in Escherichia coli K12, useful for detecting gene expression and expression patterns.

XX Claim 3; Page 1159; 2004pp; German.

XX This invention describes a novel biochip comprising probe spots, each containing many identical probes. The probes are nucleotide sequences of 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at least one includes a segment of at least 20 bases identical with, or complementary to, a segment of an open reading frame (orf) of Escherichia coli K12. The biochip is used for specific detection of gene expression in K12 and for determining the gene expression pattern, e.g. for diagnostic determination of which E. coli strains are present in the gut, and to determine the effects of e.g. growth media on gene expression. The biochip provides as comprehensive as possible detection of the K12 genome, with simultaneous analysis of many different genes with a single device, and comparison of gene expression between K12 and its mutants or other E. coli strains in a single experiment. Apart from qualitative and quantitative information about gene expression, it also allows measurements of population densities for the various strains. The use of synthetic oligonucleotides for preparation of probes allows free variation in probe length and ensures high purity (and thus selectivity, reactivity and reproducibility); also synthetic probes are generally shorter than probes prepared by polymerase chain reaction. ACD68731 to ACD81540 represent oligonucleotide probes used with the biochip described in the invention

SQ Sequence 100 BP; 24 A; 29 C; 28 G; 19 T; 0 U; 0 Other;

Query Match 17.7%; Score 21.8; DB 8; Length 100;

Best Local Similarity 52.8%; Pred. No. 3.5e+03;

Matches 47; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 9 ACGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCACCTGACAGGGGTAGACCC 68

Db 2 ACGTTTGTACAAAGGTAAAGGCGACGCCGCCCATGAAGCGCGTGCCTGCGAACCTGCG 61

QY 69 CCTCAGAGTGGCGCCCGCCATGTGCTTCAG 97

Db 62 CATGTTAATGCGCCTGTCATGGCCATGAAG 90

RESULT 13  
ACH89320  
ID ACH89320 standard; DNA; 111 BP.  
XX  
AC ACH89320;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Human genome derived single exon probe #22515.  
XX  
KW Human; probe; ss; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX  
FN US2003194704-A1.  
XX  
PD 16-OCT-2003.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PR 03-APR-2002; 2002US-00029386.  
XX  
PA (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
PI Penn SG, Rank DR, Hanzel DK;  
XX  
DR WPI; 2004-119264/12.  
XX  
PT New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
PS Claim 1; SEQ ID NO 22515; 80pp; English.  
XX  
CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridizes under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX  
SQ Sequence 111 BP; 18 A; 35 C; 30 G; 28 T; 0 U; 0 Other;  
Query Match 17.7%; Score 21.8; DB 12; Length 111;  
Best Local Similarity 70.7%; Pred. No. 3.6e+03;  
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 65 ACCCCCTCACAGTGGCCGCCCATGTGCTGTCAGTGTGGGTG 105  
Db 19 AGCCCTTCACAGAGCTTGTCTCTGTGTGTCAGCAGAGTGGCTG 59  
RESULT 14  
AAX12744  
ID AAX12744 standard; DNA; 120 BP.  
XX  
AC AAX12744;  
XX  
DT 30-MAR-1999 (first entry)  
XX  
DE Human biallelic polymorphic DNA fragment stSG3935.  
XX  
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;  
KW detection; phenotypic typing; characteristic; infection; hereditary;  
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;  
KW treatment; marker; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9820165-A2.  
XX  
PD 14-MAY-1998.  
XX  
PF 05-NOV-1997; 97WO-US020313.  
XX  
PR 06-NOV-1996; 96US-0030455P.  
XX  
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
XX  
PI Lander ES, Wang D, Hudson T;  
XX  
WPI; 1998-286974/25.  
XX  
PT New isolated nucleic acid segments from the human genome - used for  
PT determining polymorphic forms for use in e.g. forensics, paternity  
PT testing or phenotypic typing for disease.  
XX  
PS Claim 1; Page 286; 310pp; English.  
XX  
CC AAX10269-X12917 are human DNA fragments which contain biallelic  
CC polymorphic markers which have been isolated using the primers  
CC represented in AAX09121-X10268. The base occupying the polymorphic site  
CC is indicated by the appropriate IUPAC-IUB ambiguity code. These fragments  
CC can be used in methods for determining polymorphic forms in an individual  
CC for use in e.g. forensics, paternity testing or for phenotypic typing for  
CC diseases such as agammaglobulinemia, diabetes insipidus, Leach-Nyhan  
CC syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,  
CC familial hypercholesterolemia, polycystic kidney disease, hereditary  
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary  
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos  
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,  
CC autoimmune diseases, inflammation, cancer, diseases of the nervous  
CC system, infection by pathogenic microorganisms, and characteristics such  
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,  
CC endurance, fertility, and susceptibility or receptivity to particular  
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid  
CC segments can also be used to produce medicaments for the treatment or  
CC prophylaxis of such diseases  
XX  
SQ Sequence 120 BP; 25 A; 37 C; 30 G; 27 T; 0 U; 1 Other;  
Query Match 17.6%; Score 21.6; DB 2; Length 120;  
Best Local Similarity 68.2%; Pred. No. 4.2e+03;



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 21:43:24 ; Search time 307 Seconds

(without alignments)  
2302.099 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_253\_375

Perfect score: 123

Sequence: 1 accaagctacgtttcttca.....tgctgtcagccggagcccc 123

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 3219360

Minimum DB seq length: 0

Maximum DB seq length: 123

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25.8	21.0	97	9	US-09-867-701-4678
C 2	23.2	18.9	101	18	US-10-674-124A-21157
C 3	23	18.7	121	10	US-09-818-875-337
C 4	23	18.7	121	15	US-09-818-875-338
C 5	23	18.7	121	15	US-10-209-787-337
C 6	23	18.7	121	15	US-10-209-787-338
C 7	23	18.7	121	16	US-10-261-185-337
C 8	23	18.7	121	16	US-10-261-185-338
C 9	23	18.7	121	17	US-10-681-074-337
C 10	23	18.7	121	17	US-10-681-074-338
C 11	22.2	18.0	117	16	US-10-242-535A-10710
C 12	22.2	18.0	117	16	US-10-085-783A-10710

13	21.8	17.7	111	15	US-10-029-386-22515	Sequence 22515, A
C 14	21.2	17.2	121	10	US-09-818-875-209	Sequence 210, App
C 15	21.2	17.2	121	10	US-09-818-875-210	Sequence 209, App
C 16	21.2	17.2	121	15	US-10-209-787-209	Sequence 209, App
C 17	21.2	17.2	121	15	US-10-209-787-210	Sequence 210, App
C 18	21.2	17.2	121	16	US-10-261-185-209	Sequence 209, App
C 19	21.2	17.2	121	16	US-10-261-185-210	Sequence 210, App
C 20	21.2	17.2	121	17	US-10-681-074-209	Sequence 209, App
C 21	21.2	17.2	121	17	US-10-681-074-210	Sequence 210, App
C 22	21	17.1	65	10	US-09-808-975-27104	Sequence 27104, A
C 23	20.8	16.9	60	10	US-09-808-975-31791	Sequence 31791, A
C 24	20.8	16.9	60	10	US-09-808-975-31960	Sequence 31960, A
C 25	20.8	16.9	106	15	US-10-029-386-18694	Sequence 18694, A
C 26	20.8	16.9	112	15	US-10-106-698-1348	Sequence 1348, App
C 27	20.6	16.7	65	10	US-09-808-975-30206	Sequence 30206, A
C 28	20.6	16.7	80	16	US-10-384-245-1030	Sequence 1030, App
C 29	20.6	16.7	112	9	US-09-864-761-31659	Sequence 31659, A
C 30	20.6	16.7	114	18	US-10-674-124A-24132	Sequence 24132, A
C 31	20.4	16.6	65	10	US-09-808-975-25943	Sequence 25943, A
C 32	20.4	16.6	65	10	US-09-808-975-27954	Sequence 27954, A
C 33	20.4	16.6	101	18	US-10-674-124A-21552	Sequence 21552, A
C 34	20.4	16.6	119	17	US-10-437-963-88002	Sequence 88002, A
C 35	20.2	16.4	60	10	US-09-808-975-11338	Sequence 11338, A
C 36	20.2	16.4	96	15	US-10-029-386-27083	Sequence 27083, A
C 37	20.2	16.4	122	18	US-10-425-115-163501	Sequence 163501, A
C 38	20.2	16.4	123	15	US-10-029-386-19292	Sequence 19292, A
C 39	20	16.3	96	9	US-09-915-873-3	Sequence 3, Appli
C 40	20	16.3	96	18	US-10-794-615-3	Sequence 3, Appli
C 41	20	16.3	105	9	US-09-864-761-32748	Sequence 32748, A
C 42	19.8	16.1	69	11	US-09-837-306-382	Sequence 382, App
C 43	19.8	16.1	69	11	US-09-837-306-383	Sequence 383, App
C 44	19.8	16.1	69	15	US-10-045-674-298	Sequence 298, App
C 45	19.8	16.1	69	15	US-10-045-674-299	Sequence 299, App

#### ALIGNMENTS

RESULT 1  
US-09-867-701-4678/C  
; Sequence 4678, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 4678  
; LENGTH: 97  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-4678

Query Match 21.0%; Score 25.8; DB 9; Length 97;  
Best Local Similarity 56.5%; Pred. No. 76;  
Matches 48; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
Qy 39 CCAGGCTACCTGACACAGGGTAGACCCCTCACAGTGGCGGCATGTGCTGAGT 98  
Db 97 CCCAGCGCGCCCTGCTTCTTCTCCGCGCCCAACCTCCAGTCGCGCCGAGA 38  
Qy 99 GTGGTCTGCTCAGCGCGGACCCC 123  
Db 37 ACTTCGCTGCTCGCGCGACCCC 13  
RESULT 2

US-10-674-124A-21157/c  
; Sequence 21157, Application US/10674124A  
; Publication No. US20040197797A1  
; GENERAL INFORMATION:  
; APPLICANT: INOKO, Hidetoshi  
; APPLICANT: TAMIYA, Gen  
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE  
; FILE REFERENCE: ORIN-003CIP  
; CURRENT APPLICATION NUMBER: US/10/674,124A  
; PRIOR FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 10/257,511  
; PRIOR FILING DATE: 2003-03-07  
; PRIOR APPLICATION NUMBER: PCT/JP00/07621  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: JP2000-112699  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: JP2002-327516  
; PRIOR FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: JP2002-383869  
; PRIOR FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 27110  
; SEQ ID NO 21157  
; LENGTH: 101  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: DID22N\_0010585  
; FEATURE:  
; OTHER INFORMATION: Located on chromosome 14  
; FEATURE:  
; OTHER INFORMATION: Distance between a terminus base of telomere on  
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base  
; OTHER INFORMATION: sequence : 104184465  
; FEATURE:  
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of  
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and  
; OTHER INFORMATION: 5'-terminus of this base sequence : 76372  
US-10-674-124A-21157

Query Match 18.9%; Score 23.2; DB 18; Length 101;  
Best Local Similarity 56.6%; Pred. No. 5.5e+02;  
Matches 43; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
Qy 12 TTTCTTCAGACTGAGGACAGCTGGGTCCCGAGCTCCTACCTGACACAGGGGCTAGACCCCT 71  
Db 101 TCTCTCTTAATGAGCAGTCCCTCACCCTGCTTCCCTTCACACACACACACACACA 42  
Qy 72 CACAGTGGCGGCCAT 87  
Db 41 CACTGTGTGTCCAT 26

RESULT 3  
US-09-818-875-337/c  
; Sequence 337, Application US/09818875  
; Publication No. US20030051270A1  
; GENERAL INFORMATION:  
; APPLICANT: Kmiec, Eric B.  
; APPLICANT: Gamper, Howard B.  
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
; FILE REFERENCE: Napro-4  
; CURRENT APPLICATION NUMBER: US/09/818,875  
; CURRENT FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 60/192,176  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 60/192,179  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 60/208,538  
; PRIOR FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: US 60/244,989

; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 4385  
; SOFTWARE: Friedman macro Napro4  
; SEQ ID NO 337  
; LENGTH: 121  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-818-875-337

Query Match 18.7%; Score 23; DB 10; Length 121;  
Best Local Similarity 54.0%; Pred. No. 6.4e+02;  
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
Qy 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCCGAGCTCCTACCTGACACAGGGGCTAGACCCC 69  
Db 118 CGCTTAGTGTCTCCCTGGGGGAGCTCGTGTGAGGCTCCCTTTCTTGGGAGATTCTCT 59  
Qy 70 CTCACAGTGGCGGCCATGTGCTGCA 96  
Db 58 TCCTCTGTGGCGGCTCTCTCCAGGA 32

RESULT 4  
US-09-818-875-338  
; Sequence 338, Application US/09818875  
; Publication No. US20030051270A1  
; GENERAL INFORMATION:  
; APPLICANT: Kmiec, Eric B.  
; APPLICANT: Gamper, Howard B.  
; APPLICANT: Rice, Michael C.  
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
; FILE REFERENCE: Napro-4  
; CURRENT APPLICATION NUMBER: US/09/818,875  
; CURRENT FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 60/192,176  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 60/192,179  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 60/208,538  
; PRIOR FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: US 60/244,989  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 4385  
; SOFTWARE: Friedman macro Napro4  
; SEQ ID NO 338  
; LENGTH: 121  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-818-875-338

Query Match 18.7%; Score 23; DB 10; Length 121;  
Best Local Similarity 54.0%; Pred. No. 6.4e+02;  
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
Qy 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCCGAGCTCCTACCTGACACAGGGGCTAGACCCC 69  
Db 4 CGCTTAGTGTCTCCCTGGGGGAGCTCGTGTGAGGCTCCCTTTCTTGGGAGATTCTCT 63  
Qy 70 CTCACAGTGGCGGCCATGTGCTGCA 96  
Db 64 TCCTCTGTGGCGGCTCTCTCCAGGA 90

RESULT 5  
US-10-209-787-337/c  
; Sequence 337, Application US/10209787  
; Publication No. US20030217377A1  
; GENERAL INFORMATION:  
; APPLICANT: Kmiec, Eric B.  
; APPLICANT: Gamper, Howard B.  
; APPLICANT: Rice, Michael C.  
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single



;; TITLE OF INVENTION: Stranded Oligonucleotides  
;; FILE REFERENCE: Napro-4  
;; CURRENT APPLICATION NUMBER: US/10/209,787  
;; CURRENT FILING DATE: 2002-07-30  
;; PRIOR APPLICATION NUMBER: US 09/818,875  
;; PRIOR FILING DATE: 2001-03-27  
;; PRIOR APPLICATION NUMBER: US 60/192,176  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 60/192,179  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 60/208,538  
;; PRIOR FILING DATE: 2000-06-01  
;; PRIOR APPLICATION NUMBER: US 60/244,989  
;; PRIOR FILING DATE: 2000-10-30  
;; NUMBER OF SEQ ID NOS: 4385  
;; SOFTWARE: Friedman macro Napro4  
;; SEQ ID NO 337  
;; LENGTH: 121  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-209-787-337

Query Match 18.7%; Score 23; DB 15; Length 121;  
Best Local Similarity 54.0%; Pred. No. 6.4e+02;  
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
  
Qy 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCACCTGACACAGGGCTAGACCCC 69  
Db 118 CGTTAGTGTCTCCCTGGGGGAGCTCGTGTGAGGCTCCCTTTCTTTCGGGAGATTCTCT 59  
  
Qy 70 CTCACAGTGGCGGCCCATGTGCTGCA 96  
Db 58 TCCTCTGTGGCGCGGTCTCTCCACGGA 32

RESULT 6  
US-10-209-787-338  
;; Sequence 338, Application US/10209787  
;; Publication No. US20030217377A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Kmiec, Eric B.  
;; APPLICANT: Gamper, Howard B.  
;; APPLICANT: Rice, Michael C.  
;; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
;; TITLE OF INVENTION: Stranded Oligonucleotides  
;; FILE REFERENCE: Napro-4  
;; CURRENT APPLICATION NUMBER: US/10/209,787  
;; CURRENT FILING DATE: 2002-07-30  
;; PRIOR APPLICATION NUMBER: US 09/818,875  
;; PRIOR FILING DATE: 2001-03-27  
;; PRIOR APPLICATION NUMBER: US 60/192,176  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 60/192,179  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 60/208,538  
;; PRIOR FILING DATE: 2000-06-01  
;; PRIOR APPLICATION NUMBER: US 60/244,989  
;; PRIOR FILING DATE: 2000-10-30  
;; NUMBER OF SEQ ID NOS: 4385  
;; SOFTWARE: Friedman macro Napro4  
;; SEQ ID NO 338  
;; LENGTH: 121  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-209-787-338

Query Match 18.7%; Score 23; DB 15; Length 121;  
Best Local Similarity 54.0%; Pred. No. 6.4e+02;  
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
  
Qy 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCACCTGACACAGGGCTAGACCCC 69  
Db 4 CGTTAGTGTCTCCCTGGGGGAGCTCGTGTGAGGCTCCCTTTCTTTCGGGAGATTCTCT 63

Qy 70 CTCACAGTGGCGGCCCATGTGCTGCA 96  
Db 64 TCCTCTGTGGCGCGGTCTCTCCACGGA 90

RESULT 7  
US-10-261-185-337/c  
;; Sequence 337, Application US/10261185  
;; Publication No. US20040014057A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Kmiec, Eric B.  
;; APPLICANT: Gamper, Howard B.  
;; APPLICANT: Rice, Michael C.  
;; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
;; TITLE OF INVENTION: Stranded Oligonucleotides  
;; FILE REFERENCE: Napro-4CON  
;; CURRENT APPLICATION NUMBER: US/10/261,185  
;; CURRENT FILING DATE: 2002-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/09761  
;; PRIOR FILING DATE: 2001-03-27  
;; PRIOR APPLICATION NUMBER: US 60/192,176  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 60/192,179  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 60/208,538  
;; PRIOR FILING DATE: 2000-06-01  
;; PRIOR APPLICATION NUMBER: US 60/244,989  
;; PRIOR FILING DATE: 2000-10-30  
;; NUMBER OF SEQ ID NOS: 4385  
;; SOFTWARE: Friedman macro Napro4  
;; SEQ ID NO 337  
;; LENGTH: 121  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-261-185-337

Query Match 18.7%; Score 23; DB 16; Length 121;  
Best Local Similarity 54.0%; Pred. No. 6.4e+02;  
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
  
Qy 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCACCTGACACAGGGCTAGACCCC 69  
Db 118 CGTTAGTGTCTCCCTGGGGGAGCTCGTGTGAGGCTCCCTTTCTTTCGGGAGATTCTCT 59  
  
Qy 70 CTCACAGTGGCGGCCCATGTGCTGCA 96  
Db 58 TCCTCTGTGGCGCGGTCTCTCCACGGA 32

RESULT 8  
US-10-261-185-338  
;; Sequence 338, Application US/10261185  
;; Publication No. US20040014057A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Kmiec, Eric B.  
;; APPLICANT: Gamper, Howard B.  
;; APPLICANT: Rice, Michael C.  
;; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
;; TITLE OF INVENTION: Stranded Oligonucleotides  
;; FILE REFERENCE: Napro-4CON  
;; CURRENT APPLICATION NUMBER: US/10/261,185  
;; CURRENT FILING DATE: 2002-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/09761  
;; PRIOR FILING DATE: 2001-03-27  
;; PRIOR APPLICATION NUMBER: US 60/192,176  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 60/192,179  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 60/208,538  
;; PRIOR FILING DATE: 2000-06-01  
;; PRIOR APPLICATION NUMBER: US 60/244,989  
;; PRIOR FILING DATE: 2000-10-30

```

; NUMBER OF SEQ ID NOS: 4385
; SOFTWARE: Friedman macro Napro4
; SEQ ID NO 338
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-261-185-338

```

Query Match 18.7%; Score 23; DB 16; Length 121;  
Best Local Similarity 54.0%; Pred. No. 6.4e+02;  
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

70 CTCACAGTGGCGCCGCAATGTGCCTGCA 96  
| | | | |  
64 TCCTCTGTGGCGCCGGTCTCTCCAGGA 90

```

RESULT 9
US-10-681-074-337/c
; Sequence 337, Application US/10681074
; Publication No. US20040175722A1
; GENERAL INFORMATION:
; APPLICANT: KMEIC, ERIC B.
; APPLICANT: VAN BRABANT, ANJA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR REDUCING SCREENING IN
; TITLE OF INVENTION: OLIGONUCLEOTIDE-DIRECTED NUCLEIC ACID SEQUENCE ALTERATION
; FILE REFERENCE: NaPro-18 US
; CURRENT APPLICATION NUMBER: US/10/681.074
; CURRENT FILING DATE: 2003-10-07
; PRIOR APPLICATION NUMBER: US 60/453,360
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: US 60/416,983
; PRIOR FILING DATE: 2002-10-07
; NUMBER OF SEQ ID NOS: 4375
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 337
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-681-074-337

```

	Query Match	18.7%	Score 23	DB 17	Length 121
	Best Local Similarity	54.0%	Pred. No. 6.4e+02		
	Matches 47	Conservative 0	Mismatches 40	Indels 0	Gaps 0
QY	10	CGTTTCCTTCAGACTGAGGACAGCTGGGTCCCGAGCTCACTGTACACAGGGCTAGACCCC	69		
Db	118	CGCTTAGTGTCTCCCTGGGGGAGCTCGTGTGTAGGCTCCCCCTTCTTGGGGAGATTCTCT	59		
QY	70	CTCAGATGCGCGCCCATGTGCTGCA	96		
Db	58	TCCTCTGTGCGCGGTCTCTCCACGGA	32		

RESULT 10  
US-10-681-074-338  
; Sequence 338, Application US/10681074  
; Publication No. US20040175722A1  
; GENERAL INFORMATION:  
; APPLICANT: KMEC, ERIC B.  
; APPLICANT: VAN BRABANT, ANJA  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR REDUCING SCREENING IN  
; TITLE OF INVENTION: OLIGONUCLEOTIDE-DIRECTED NUCLEIC ACID SEQUENCE ALTERATION  
; FILE REFERENCE: NaPro-18 US  
; CURRENT APPLICATION NUMBER: US/10/681,074  
; CURRENT FILING DATE: 2003-10-07  
; PRIOR APPLICATION NUMBER: US 60/453,360  
; PRIOR FILING DATE: 2003-03-07  
; PRIOR APPLICATION NUMBER: US 60/416,983

```

; PRIOR FILING DATE: 2002-10-07
; NUMBER OF SEQ ID NOS: 4375
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 338
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-681-074-338

```

Query Match 18.7%; Score 23; DB 17; Length 121;  
Best Local Similarity 54.0%; Pred. No. 6.4e+02;  
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 70 CTCACAGTGGCGCGCCATGTGCCTGCA 96  
Dd 64 TCCTCTGTGCGCGGTCTCTCCAGGA 90

RESULT 11  
 US-10-242-535A-10710/c  
 ; Sequence 10710, Application US/10242535A  
 ; Publication NO. US20040013663A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ChondroGene Inc.  
 ; APPLICANT: Liew, C.C.  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
 ; FILE REFERENCE: 4231/2005  
 ; CURRENT APPLICATION NUMBER: US/10/242,535A  
 ; CURRENT FILING DATE: 2002-09-12  
 ; PRIOR APPLICATION NUMBER: US 10/085,783  
 ; PRIOR FILING DATE: 2002-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/305,340  
 ; PRIOR FILING DATE: 2001-07-13  
 ; PRIOR APPLICATION NUMBER: US 60/275,017  
 ; PRIOR FILING DATE: 2001-03-12  
 ; PRIOR APPLICATION NUMBER: US 60/271,955  
 ; PRIOR FILING DATE: 2001-02-28  
 ; NUMBER OF SEQ ID NOS: 58994  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 10710  
 ; LENGTH: 117

;; PRIOR FILING DATE: 2003-03-07  
;; PRIOR APPLICATION NUMBER: US 60/453,360  
;; PRIOR APPLICATION NUMBER: US 60/416,983



[illegible]

**RESULT 15**

```

US-09-818-875-210
; Sequence 210, Application US/09818875
; Publication No. US20030051270A1
; GENERAL INFORMATION:
; APPLICANT: Kmtec, Eric B.
; APPLICANT: Gamper, Howard B.
; APPLICANT: Rice, Michael C.
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
; TITLE OF INVENTION: Stranded Oligonucleotides
; FILE REFERENCE: Napro-4
; CURRENT APPLICATION NUMBER: US/09/818,875
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/192,176
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/192,179
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/208,538
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/244,989
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 4385
; SOFTWARE: Friedman macro Napro4
; SEQ ID NO 210
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-818-875-210

```

	Query Match	17.28;	Score 21.2;	DB 10;	Length 121;
	Best Local Similarity	55.44;	Pred. No. 2.5e+03;		
	Matches	41;	Conservative	0;	Mismatches 33;
					Indels 0;
					Gaps 0;
Qy	23	CTGAGCAGCTGGGTCCCGCTCACTGACACAGGGCTAGACCCCTCACAGTGGCC	82		
Db	1	CTGGGGCAGCTCGGTGAGGCTCCCTTTCTTGCAGATTCTCTCTGTGGCC	60		
Qy	83	GCCATGTGCTGCA	96.		
Db	61	GGTCTCTCCAGGA	74		

Search completed: January 31, 2005, 23:05:45  
Job time : 308 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 19:01:47 ; Search time 2458 Seconds  
(without alignments)

4503.816 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_1\_273

Perfect score: 273

Sequence: 1 atggagcactaccggaagc.....ccaagtacgttccttcag 273

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 44308572 seqs, 20275418765 residues

Total number of hits satisfying chosen parameters: 55031126

Minimum DB seq length: 0

Maximum DB seq length: 273

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:\*

- 1: /cgn2\_6/ptodata/1/pna/PCTUS1 COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pna/PCTUS2 COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pna/US06 COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pna/US07 COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pna/US08 COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pna/US081 COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pna/US082 COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pna/US083 COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pna/US084 COMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pna/US085 COMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pna/US086 COMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pna/US087 COMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pna/US088 COMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pna/US089 COMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pna/US090 COMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pna/US091 COMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pna/US092A COMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pna/US092B COMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pna/US093A COMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pna/US093B COMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pna/US094 COMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pna/US095A COMB.seq:\*
- 23: /cgn2\_6/ptodata/1/pna/US095B COMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pna/US095C COMB.seq:\*
- 25: /cgn2\_6/ptodata/1/pna/US095D COMB.seq:\*
- 26: /cgn2\_6/ptodata/1/pna/US096A COMB.seq:\*
- 27: /cgn2\_6/ptodata/1/pna/US096B COMB.seq:\*
- 28: /cgn2\_6/ptodata/1/pna/US096C COMB.seq:\*
- 29: /cgn2\_6/ptodata/1/pna/US096D COMB.seq:\*
- 30: /cgn2\_6/ptodata/1/pna/US096E COMB.seq:\*
- 31: /cgn2\_6/ptodata/1/pna/US097A COMB.seq:\*
- 32: /cgn2\_6/ptodata/1/pna/US097B COMB.seq:\*
- 33: /cgn2\_6/ptodata/1/pna/US097C COMB.seq:\*
- 34: /cgn2\_6/ptodata/1/pna/US098A COMB.seq:\*
- 35: /cgn2\_6/ptodata/1/pna/US098B COMB.seq:\*
- 36: /cgn2\_6/ptodata/1/pna/US098C COMB.seq:\*
- 37: /cgn2\_6/ptodata/1/pna/US098D COMB.seq:\*
- 38: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*
- 39: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*
- 40: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*
- 41: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*
- 42: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*
- 43: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

- 44: /cgn2\_6/ptodata/1/pna/US099G COMB.seq:\*
- 45: /cgn2\_6/ptodata/1/pna/US100A COMB.seq:\*
- 46: /cgn2\_6/ptodata/1/pna/US100B COMB.seq:\*
- 47: /cgn2\_6/ptodata/1/pna/US101A COMB.seq:\*
- 48: /cgn2\_6/ptodata/1/pna/US101B COMB.seq:\*
- 49: /cgn2\_6/ptodata/1/pna/US102A COMB.seq:\*
- 50: /cgn2\_6/ptodata/1/pna/US102B COMB.seq:\*
- 51: /cgn2\_6/ptodata/1/pna/US103A COMB.seq:\*
- 52: /cgn2\_6/ptodata/1/pna/US103B COMB.seq:\*
- 53: /cgn2\_6/ptodata/1/pna/US104A COMB.seq:\*
- 54: /cgn2\_6/ptodata/1/pna/US104B COMB.seq:\*
- 55: /cgn2\_6/ptodata/1/pna/US105A COMB.seq:\*
- 56: /cgn2\_6/ptodata/1/pna/US105B COMB.seq:\*
- 57: /cgn2\_6/ptodata/1/pna/US106A COMB.seq:\*
- 58: /cgn2\_6/ptodata/1/pna/US107A COMB.seq:\*
- 59: /cgn2\_6/ptodata/1/pna/US107B COMB.seq:\*
- 60: /cgn2\_6/ptodata/1/pna/US107C COMB.seq:\*
- 61: /cgn2\_6/ptodata/1/pna/US107D COMB.seq:\*
- 62: /cgn2\_6/ptodata/1/pna/US108A COMB.seq:\*
- 63: /cgn2\_6/ptodata/1/pna/US108B COMB.seq:\*
- 64: /cgn2\_6/ptodata/1/pna/US109A COMB.seq:\*
- 65: /cgn2\_6/ptodata/1/pna/US109B COMB.seq:\*
- 66: /cgn2\_6/ptodata/1/pna/US109C COMB.seq:\*
- 67: /cgn2\_6/ptodata/1/pna/US109D COMB.seq:\*
- 68: /cgn2\_6/ptodata/1/pna/US109E COMB.seq:\*
- 69: /cgn2\_6/ptodata/1/pna/US109F COMB.seq:\*
- 70: /cgn2\_6/ptodata/1/pna/US109G COMB.seq:\*
- 71: /cgn2\_6/ptodata/1/pna/US109H COMB.seq:\*
- 72: /cgn2\_6/ptodata/1/pna/US109I COMB.seq:\*
- 73: /cgn2\_6/ptodata/1/pna/US109J COMB.seq:\*
- 74: /cgn2\_6/ptodata/1/pna/US109K COMB.seq:\*
- 75: /cgn2\_6/ptodata/1/pna/US109L COMB.seq:\*
- 76: /cgn2\_6/ptodata/1/pna/US109M COMB.seq:\*
- 77: /cgn2\_6/ptodata/1/pna/US109N COMB.seq:\*
- 78: /cgn2\_6/ptodata/1/pna/US109O COMB.seq:\*
- 79: /cgn2\_6/ptodata/1/pna/US109P COMB.seq:\*
- 80: /cgn2\_6/ptodata/1/pna/US109Q COMB.seq:\*
- 81: /cgn2\_6/ptodata/1/pna/US109R COMB.seq:\*
- 82: /cgn2\_6/ptodata/1/pna/US109S COMB.seq:\*
- 83: /cgn2\_6/ptodata/1/pna/US109T COMB.seq:\*
- 84: /cgn2\_6/ptodata/1/pna/US109U COMB.seq:\*
- 85: /cgn2\_6/ptodata/1/pna/US109V COMB.seq:\*
- 86: /cgn2\_6/ptodata/1/pna/US109W COMB.seq:\*
- 87: /cgn2\_6/ptodata/1/pna/US109X COMB.seq:\*
- 88: /cgn2\_6/ptodata/1/pna/US109Y COMB.seq:\*
- 89: /cgn2\_6/ptodata/1/pna/US109Z COMB.seq:\*
- 90: /cgn2\_6/ptodata/1/pna/US109A COMB.seq:\*
- 91: /cgn2\_6/ptodata/1/pna/US109B COMB.seq:\*
- 92: /cgn2\_6/ptodata/1/pna/US109C COMB.seq:\*
- 93: /cgn2\_6/ptodata/1/pna/US109D COMB.seq:\*
- 94: /cgn2\_6/ptodata/1/pna/US109E COMB.seq:\*
- 95: /cgn2\_6/ptodata/1/pna/US109F COMB.seq:\*
- 96: /cgn2\_6/ptodata/1/pna/US109G COMB.seq:\*
- 97: /cgn2\_6/ptodata/1/pna/US109H COMB.seq:\*
- 98: /cgn2\_6/ptodata/1/pna/US109I COMB.seq:\*
- 99: /cgn2\_6/ptodata/1/pna/US109J COMB.seq:\*
- 100: /cgn2\_6/ptodata/1/pna/US109K COMB.seq:\*
- 101: /cgn2\_6/ptodata/1/pna/US109L COMB.seq:\*
- 102: /cgn2\_6/ptodata/1/pna/US109M COMB.seq:\*
- 103: /cgn2\_6/ptodata/1/pna/US109N COMB.seq:\*
- 104: /cgn2\_6/ptodata/1/pna/US109O COMB.seq:\*
- 105: /cgn2\_6/ptodata/1/pna/US109P COMB.seq:\*
- 106: /cgn2\_6/ptodata/1/pna/US109Q COMB.seq:\*
- 107: /cgn2\_6/ptodata/1/pna/US109R COMB.seq:\*
- 108: /cgn2\_6/ptodata/1/pna/US109S COMB.seq:\*
- 109: /cgn2\_6/ptodata/1/pna/US109T COMB.seq:\*
- 110: /cgn2\_6/ptodata/1/pna/US109U COMB.seq:\*
- 111: /cgn2\_6/ptodata/1/pna/US109V COMB.seq:\*
- 112: /cgn2\_6/ptodata/1/pna/US109W COMB.seq:\*
- 113: /cgn2\_6/ptodata/1/pna/US109X COMB.seq:\*
- 114: /cgn2\_6/ptodata/1/pna/US109Y COMB.seq:\*
- 115: /cgn2\_6/ptodata/1/pna/US109Z COMB.seq:\*
- 116: /cgn2\_6/ptodata/1/pna/US109A COMB.seq:\*

117: /cgn2\_6/ptodata/1/pna/us6051\_COMB.seq.\*  
118: /cgn2\_6/ptodata/1/pna/us6052\_COMB.seq.\*  
119: /cgn2\_6/ptodata/1/pna/us6053\_COMB.seq.\*  
120: /cgn2\_6/ptodata/1/pna/us6054\_COMB.seq.\*  
121: /cgn2\_6/ptodata/1/pna/us6055\_COMB.seq.\*  
122: /cgn2\_6/ptodata/1/pna/us6056\_COMB.seq.\*  
123: /cgn2\_6/ptodata/1/pna/us6057\_COMB.seq.\*  
124: /cgn2\_6/ptodata/1/pna/us6058\_COMB.seq.\*  
125: /cgn2\_6/ptodata/1/pna/us6059\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
1	226	82.8	264	13	US-08-856-624-2456	Sequence 2456, Ap
2	226	82.8	264	24	US-09-540-766-8863	Sequence 8863, Ap
3	218	79.9	271	14	US-08-941-869A-2841	Sequence 2841, Ap
4	218	79.9	271	24	US-09-540-208-35269	Sequence 35269, A
5	212	77.7	245	9	US-08-438-571A-2389	Sequence 2389, Ap
6	212	77.7	245	24	US-09-540-233D-126181	Sequence 126181, Ap
7	206	75.5	260	14	US-08-959-395-768	Sequence 768, App
8	206	75.5	260	23	US-09-539-334-6888	Sequence 6888, Ap
9	196	71.8	262	13	US-08-812-505-1415	Sequence 1415, Ap
10	196	71.8	262	14	US-08-978-620-1415	Sequence 1415, Ap
11	196	71.8	262	24	US-09-540-229-45100	Sequence 45100, A
12	196	71.8	262	66	US-06-013-529-1415	Sequence 1415, Ap
13	193	70.7	250	14	US-08-951-197-5067	Sequence 5067, Ap
14	193	70.7	250	23	US-09-539-806-30498	Sequence 30498, A
15	193	70.7	250	23	US-09-539-806B-30498	Sequence 30498, A
16	193	70.7	250	67	US-06-027-249-1037	Sequence 1037, Ap
17	192	70.3	250	23	US-09-539-800-16108	Sequence 16108, A
18	192	70.3	250	23	US-09-539-800B-16108	Sequence 16108, A
19	192	70.3	250	23	US-09-539-800C-16108	Sequence 16108, A
20	190	69.6	244	15	US-09-014-441-1305	Sequence 1305, Ap
21	190	69.6	244	24	US-09-540-212A-30093	Sequence 30093, A
22	190	69.6	244	68	US-06-034-841-1305	Sequence 1305, Ap
23	189.6	69.5	264	24	US-09-540-213-18789	Sequence 18789, A
24	185	67.8	239	14	US-08-901-902-73	Sequence 73, Appl
25	185	67.8	239	23	US-09-539-806-30137	Sequence 30137, A
26	185	67.8	239	23	US-09-539-806B-30137	Sequence 30137, A
27	185	67.8	239	67	US-06-023-379-73	Sequence 73, Appl
28	184	67.4	233	13	US-08-879-863-3710	Sequence 3710, Ap
29	184	67.4	233	13	US-08-879-863A-3710	Sequence 3710, Ap
30	184	67.4	233	24	US-09-540-213-12679	Sequence 12679, A
31	183.4	67.2	259	18	US-09-293-657-1609	Sequence 1609, Ap
32	183.4	67.2	259	24	US-09-540-233D-76641	Sequence 76641, Ap
33	183.4	67.2	259	73	US-06-082-035-1609	Sequence 1609, Ap
34	183	67.0	238	23	US-09-539-334-28833	Sequence 28833, A
35	181	66.3	238	24	US-09-540-229-113556	Sequence 113556, A
36	169	61.9	227	9	US-08-413-150-22	Sequence 22, Appl
37	169	61.9	227	24	US-09-540-499-10618	Sequence 10618, A
38	169	61.9	227	24	US-09-540-499B-10618	Sequence 10618, A
39	166	60.8	224	23	US-09-539-800-15324	Sequence 15324, A
40	166	60.8	224	23	US-09-539-800B-15324	Sequence 15324, A
41	166	60.8	224	23	US-09-539-800C-15324	Sequence 15324, A
42	166	60.8	224	14	US-08-923-902-2187	Sequence 2187, Ap
43	166	60.8	224	24	US-09-540-208-7903	Sequence 7903, Ap
44	157	57.5	210	13	US-08-801-504-2052	Sequence 2052, Ap
45	157	57.5	210	24	US-09-540-499-16308	Sequence 16308, A

## ALIGNMENTS

RESULT 1  
US-08-856-624-2456  
; Sequence 2456, Application US/08856624  
; GENERAL INFORMATION:

APPLICANT: Gooding, Douglas H.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Akerblom, Ingrid E.  
APPLICANT: Deleane, Angelo M.  
APPLICANT: Naughton, Rebecca E.  
APPLICANT: Klingler, Tod M.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
TITLE OF INVENTION: SIGMOID ADENOCARCINOMA  
NUMBER OF SEQUENCES: 3070  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/856,624  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/017,190  
FILING DATE: MAY 13, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,100  
FILING DATE: SEPTEMBER 25, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: CERRONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PD-0174 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 855-0555  
TELEFAX: (415) 845-4166  
INFORMATION FOR SEQ ID NO: 2456:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: 1929519  
US-08-856-624-2456

Query Match 82.8%; Score 226; DB 13; Length 264;  
Best Local Similarity 100.0%; Pred. No. 6.4e-51;  
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CCCAATGCCAGCTACCTCTGATACCTTGAGATGCGGTCGAGATGCGAGCAAAAT 107  
Db 1 CCCAATGCCAGCTACCTCTGATACCTTGAGATGCGGTCGAGATGCGAGCAAAAT 60  
QY 108 TCGCAACCTGCTGGGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTGTCTCGCATGTAGT 167  
Db 61 TCGCAACCTGCTGGGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTGTCTCGCATGTAGT 120  
QY 168 GTTCTCAGGTTCTGGCAGGGCTGACAGAAAGGCTGTAGTGTGCGCTGAGATTGCAACGC 227  
Db 121 GTTCTCAGGTTCTGGCAGGGCTGACAGAAAGGCTGTAGTGTGCGCTGAGATTGCAACGC 180  
QY 228 GCGGTCGCCAGCTGACACAGCTCACCAGCTACGTTTCCTTCAG 273  
Db 181 GCGGTCGCCAGCTGACACAGCTCACCAGCTACGTTTCCTTCAG 226

RESULT 2

```
US-09-540-766-8863
; Sequence 8863, Application US/09540766
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
; FILE REFERENCE: PD-1024 CIP
; CURRENT APPLICATION NUMBER: US/09/540,766
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION DATA removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 77960
; SOFTWARE: PERL Program
; SEQ ID NO 8863
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: incyte ID No: hu00958497
US-09-540-766-8863

Query Match      82.8%; Score 226; DB 24; Length 264;
Best Local Similarity 100.0%; Pred. No. 6.4e-51;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CCCAATGCCAGCTACCTCTGTATACCTTGAGATCGGGTCGAGATGCGCAAAAT 107
Db 1 CCCAATGCCAGCTACCTCTGTATACCTTGAGATCGGGTCGAGATGCGCAAAAT 60

QY 108 TCGCAACCTGCTGGGTTGGCTCGGTCGGTTCGAGGCGGCGAGTCTCGGCATGTAGT 167
Db 61 TCGCAACCTGCTGGGTTGGCTCGGTCGGTTCGAGGCGGCGAGTCTCGGCATGTAGT 120

QY 168 GTTCTCAGGTTCTGGCAGGCTGCGAGGAGGCTGTGAGTTCGCTGAGATTTGTCAGGCG 227
Db 121 GTTCTCAGGTTCTGGCAGGCTGCGAGGAGGCTGTGAGTTCGCTGAGATTTGTCAGGCG 180

QY 228 GCGGTCGCCAGCTGACACAGCTCACAAGCTACGTTTCCTTCAG 273
Db 181 GCGGTCGCCAGCTGACACAGCTCACAAGCTACGTTTCCTTCAG 226

RESULT 3
US-08-941-869A-2841
; Sequence 2841, Application US/08941869A
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 5486
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
```

```
US-08-941-869A-2841
; Sequence 2841, Application US/08941869A
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 5486
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

US-09-540-208-35269
; Sequence 35269, Application US/09540208
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE
; FILE REFERENCE: PD-1029 CIP
; CURRENT APPLICATION NUMBER: US/09/540,208
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION DATA removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 70811
; SOFTWARE: PERL Program
; SEQ ID NO 35269
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

Query Match      79.9%; Score 218; DB 14; Length 271;
Best Local Similarity 100.0%; Pred. No. 9.7e-49;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCAGCTACCGAAAGCTGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCGAG 60
Db 54 ATGAGCAGCTACCGAAAGCTGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCGAG 113

QY 61 CTACCTCTCATACCTTGAGATCGGGTCGAGATGCGAGATGCGCAAAATTCGCAACTGCTG 120
Db 114 CTACCTCTCATACCTTGAGATCGGGTCGAGATGCGAGATGCGCAAAATTCGCAACTGCTG 173

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTCTCGCATGTAGTGTCTCAGGTTCT 180
Db 174 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTCTCGCATGTAGTGTCTCAGGTTCT 233

QY 181 GGCAGGCTCCAGAAAGCTGTGAGTGGCTGAGAT 218
Db 234 GGCAGGCTCCAGAAAGCTGTGAGTGGCTGAGAT 271

RESULT 4
US-09-540-208-35269
; Sequence 35269, Application US/09540208
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE
; FILE REFERENCE: PD-1029 CIP
; CURRENT APPLICATION NUMBER: US/09/540,208
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION DATA removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 70811
; SOFTWARE: PERL Program
; SEQ ID NO 35269
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: hu00317250  
US-09-540-208-35269

Query Match 79.9%; Score 218; DB 24; Length 271;  
Best Local Similarity 100.0%; Pred. No. 9.7e-49;  
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60  
DB 54 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 113  
QY 61 CTACCTCTGTATACCTTTGAGATGCGGGTCCAGATGGCAGCAAAATTCGCAACTGCTG 120  
DB 114 CTACCTCTGTATACCTTTGAGATGCGGGTCCAGATGGCAGCAAAATTCGCAACTGCTG 173  
QY 121 GGGTTGGCTCTGGGTTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
DB 174 GGGTTGGCTCTGGGTTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 233  
QY 181 GGCAGGGCTGCAGGAAAGCTGTCTAGCTGCGCTGAGAT 218  
DB 234 GGCAGGGCTGCAGGAAAGCTGTCTAGCTGCGCTGAGAT 271

## RESULT 5

US-08-438-571A-2389  
; Sequence 2389, Application US/08438571A  
; GENERAL INFORMATION:  
; APPLICANT: Deleageane, Angelo M.  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Levine, Wendy B.  
; APPLICANT: Wilde, Craig G.  
; APPLICANT: Freitas, Nicole  
; APPLICANT: Jernigan-Kelleher, Colleen  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Scott, Randall W.  
; APPLICANT: Bills, Pamela Kay  
; APPLICANT: Pham, Mino Thu  
; APPLICANT: Altus, Christina M.  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Akerblom, Ingrid Erika  
; TITLE OF INVENTION: POLYNUCLEOTIDES DERIVED FROM THP-1 CELLS  
; NUMBER OF SEQUENCES: 5094  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/438,571A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CERRONE, MICHAEL C., Ph.D.  
; REGISTRATION NUMBER: 39132  
; REFERENCE/DOCKET NUMBER: PD-001-5 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 855-0555  
; TELEFAX: (415) 845-4166  
; INFORMATION FOR SEQ ID NO: 2389:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 245 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: CDNA  
; IMMEDIATE SOURCE:  
; CLONE: P-033488  
US-08-438-571A-2389

Query Match 77.7%; Score 212; DB 9; Length 245;  
Best Local Similarity 99.1%; Pred. No. 4.1e-47;  
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60  
DB 25 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 84  
QY 61 CTACCTCTGTATACCTTTGAGATGCGGGTCCAGATGGCAGCAAAATTCGCAACTGCTG 120  
DB 85 CTACCTCTGTATACCTTTANATTCGCGGTCCGAGATGGCAGCAAAATTCGCAACTGCTG 144  
QY 121 GGGTTGGCTCTGGGTTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 180  
DB 145 GGGTTGGCTCTGGGTTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 204  
QY 181 GGCAGGGCTGCAGGAAAGCTGTCTAGCTGCGCTGAG 216  
DB 205 GGCAGGGCTGCAGGAAAGCTGTCTAGCTGCGCTGAG 240

## RESULT 6

US-09-540-233D-126181  
; Sequence 126181, Application US/09540233D  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Deleageane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF HEMIC AND IMMUNE SYSTEM TISSUE  
; FILE REFERENCE: PD-1030 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,233D  
; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/008,119  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 8/100,523  
; PRIOR FILING DATE: 1993-08-03  
; PRIOR APPLICATION NUMBER: 07/977,780  
; PRIOR FILING DATE: 1992-11-19  
; PRIOR APPLICATION NUMBER: 07/916,491  
; PRIOR FILING DATE: 1992-07-17  
; PRIOR APPLICATION NUMBER: 08/438,571  
; PRIOR FILING DATE: 1995-05-10  
; PRIOR APPLICATION NUMBER: 08/282,991  
; PRIOR FILING DATE: 1994-07-28  
; PRIOR APPLICATION NUMBER: 08/196,364  
; PRIOR FILING DATE: 1994-02-14  
; PRIOR APPLICATION NUMBER: 08/504,732  
; PRIOR FILING DATE: 1995-07-20  
; PRIOR APPLICATION NUMBER: 08/179,873  
; PRIOR FILING DATE: 1994-01-11  
; PRIOR APPLICATION NUMBER: 08/435,761  
; PRIOR FILING DATE: 1995-05-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 129340  
; SOFTWARE: PERL Program  
; SEQ ID NO 126181  
; LENGTH: 245  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: hu00448643  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 103, 105, 178, 231



OTHER INFORMATION: a, t, c, g, or other  
US-09-540-233D-126181

Query Match 77.7%; Score 212; DB 24; Length 245;  
Best Local Similarity 98.1%; Pred. No. 4.1e-47;  
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 25 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 84

Qy 61 CTACTCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 120  
Db 85 CTACTCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 144

Qy 121 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 180  
Db 145 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 204

Qy 181 GGCAGGGCTGCAGGAAAGGCTGTGAGCTGGCGCTGAG 216  
Db 205 GGCAGGGCTGCAGGAAAGGCTGTGAGCTGGCGCTGAG 240

RESULT 7  
US-08-959-395-768  
; Sequence 768, Application US/08959395  
; GENERAL INFORMATION:  
; APPLICANT: Gooding, Douglas H.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Akerblom, Ingrid E.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Klingler, Tod M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
; TITLE OF INVENTION: HUMAN PANCREATIC ISLET CELLS  
; NUMBER OF SEQUENCES: 5789  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/959,395  
; FILING DATE: HEREWITH  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/030,755  
; FILING DATE: OCTOBER 28, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/033,551  
; FILING DATE: DECEMBER 20, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CERRONE, MICHAEL C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PD-0259 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 768:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 260 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: 2071192H2  
US-08-959-395-768

Query Match 75.5%; Score 206; DB 14; Length 260;  
Best Local Similarity 100.0%; Pred. No. 1.8e-45;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 49 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 108

Qy 61 CTACTCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 120  
Db 109 CTACTCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 168

Qy 121 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 180  
Db 169 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 228

Qy 181 GGCAGGGCTGCAGGAAAGGCTGTGAG 206  
Db 229 GGCAGGGCTGCAGGAAAGGCTGTGAG 254

RESULT 8  
US-09-539-334-6888  
; Sequence 6888, Application US/09539334  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF ENDOCRINE SYSTEM TISSUE  
; FILE REFERENCE: PD-1026 CIP  
; CURRENT APPLICATION NUMBER: US/09/539,334  
; CURRENT FILING DATE: 2000-03-30  
; "Prior application data removed - refer to PALM or file wrapper"  
; NUMBER OF SEQ ID NOS: 38381  
; SOFTWARE: PERL Program  
; SEQ ID NO 6888  
; LENGTH: 260  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: hu00887627  
US-09-539-334-6888

Query Match 75.5%; Score 206; DB 23; Length 260;  
Best Local Similarity 100.0%; Pred. No. 1.8e-45;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60  
Db 49 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 108

Qy 61 CTACTCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 120  
Db 109 CTACTCTCTGATACCTTGTAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG 168

Qy 121 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 180  
Db 169 GGGTTGGCTCTGGGTCGGTTGGAGGCGGCGAGTCTCGGATGTAGTGTCTCAGGTTCT 228

Qy 181 GGCAGGGCTGCAGGAAAGGCTGTGAG 206  
Db 229 GGCAGGGCTGCAGGAAAGGCTGTGAG 254

RESULT 9  
US-08-812-505-1415  
; Sequence 1415, Application US/08812505  
; GENERAL INFORMATION:  
; APPLICANT: Gooding, Douglas H.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Akerblom, Ingrid E.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Klingler, Tod M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
; ASTROCYTOMA  
; NUMBER OF SEQUENCES: 3896  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/812,505  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/013,529  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LUTHER, BARBARA J.  
; REGISTRATION NUMBER: 33954  
; REFERENCE/DOCKET NUMBER: PD-0126P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 855-0555  
; TELEFAX: (415) 852-0195  
; INFORMATION FOR SEQ ID NO: 1415:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 262 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; IMMEDIATE SOURCE:  
; CLONE: 862319  
US-08-812-505-1415  
Query Match 71.8%; Score 196; DB 13; Length 262;  
Best Local Similarity 97.5%; Pred. No. 9.5e-43;  
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCGAG 60  
Db 60 ATGGAGCACTACCGGANAGCTGGCTCTNTAGAGCTCCAGCGCTTCCCAATGCCCGAG 119  
QY 61 CTACCTCTGATACCTTANATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 120 CTACCTCTGATACCTTANATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 179  
QY 121 GGGTTGGCTCTGGGTGGTGGAGGGCGGAGTCTCGGATGTAGTGTCTCAGTTCT 180  
Db 180 GGGTTGGCTCTGGGTGGTGGAGGGCGGAGTCTCGGATGTAGTGTCTCAGTTCT 239  
QY 181 GGCAGGCTGCAGGAAGGCT 201  
Db 240 GGCAGGCTGCAGGAAGGCT 260

RESULT 10  
US-08-978-620-1415  
; Sequence 1415, Application US/08978620  
; GENERAL INFORMATION:  
; APPLICANT: Gooding, Douglas H.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Akerblom, Ingrid E.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Klingler, Tod M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
; HUMAN BRAIN  
; NUMBER OF SEQUENCES: 7797  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,620  
; FILING DATE: HERewith  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/032,151  
; FILING DATE: DECEMBER 6, 1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/812,505  
; FILING DATE: MARCH 7, 1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LUTHER, BARBARA J.  
; REGISTRATION NUMBER: 33954  
; REFERENCE/DOCKET NUMBER: PD-0126-1 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 852-0195  
; INFORMATION FOR SEQ ID NO: 1415:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 262 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; IMMEDIATE SOURCE:  
; CLONE: 862319  
US-08-978-620-1415  
Query Match 71.8%; Score 196; DB 14; Length 262;  
Best Local Similarity 97.5%; Pred. No. 9.5e-43;  
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCGAG 60  
Db 60 ATGGAGCACTACCGGANAGCTGGCTCTNTAGAGCTCCAGCGCTTCCCAATGCCCGAG 119  
QY 61 CTACCTCTGATACCTTANATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db 120 CTACCTCTGATACCTTANATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 179  
QY 121 GGGTTGGCTCTGGGTGGTGGAGGGCGGAGTCTCGGATGTAGTGTCTCAGTTCT 180  
Db 180 GGGTTGGCTCTGGGTGGTGGAGGGCGGAGTCTCGGATGTAGTGTCTCAGTTCT 239

QY 181 GGCAGGGCTGCAGAAAGGCT 201  
|||||  
Db 240 GGCAGGGCTGCAGAAAGGCT 260

## RESULT 11

US-09-540-229-45100  
; Sequence 45100, Application US/09540229  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Deleane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS  
; FILE REFERENCE: PD-1033 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,229  
; CURRENT FILING DATE: 2000-03-31  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 193582  
; SOFTWARE: PERL Program  
; SEQ ID NO 45100  
; LENGTH: 262  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: hu00691274  
; NAME/KEY: unsure  
; LOCATION: 38, 76, 87, 138, 140, 168  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-540-229-45100

Query Match 71.8%; Score 196; DB 24; Length 262;

Best Local Similarity 97.5%; Pred. No. 9.5e-43;

Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGAGCAGCTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCTTCCCAATGCCCCAG 60  
Db 60 ATGAGCAGCTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCTTCCCAATGCCCCAG 119  
QY 61 CTACCTCTGATACCTTCCGAGTCCGAGTGGCAGCAAAATTCGCAACCTGCTG 120  
Db 120 CTACCTCTGATACCTTCCGAGTCCGAGTGGCAGCAAAATTCGCAACCTGCTG 179  
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTCTCGGATGTAGTTCTCAGGTTCT 180  
Db 180 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTCTCGGATGTAGTTCTCAGGTTCT 239  
QY 181 GGCAGGGCTGCAGAAAGGCT 201  
Db 240 GGCAGGGCTGCAGAAAGGCT 260

## RESULT 12

US-60-013-529-1415  
; Sequence 1415, Application US/60013529  
; GENERAL INFORMATION:  
; APPLICANT: Gooding, Douglas H.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Akerblom, Ingrid E.  
; APPLICANT: Deleane, Angelo M.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Klingler, Tod M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
; NUMBER OF SEQUENCES: 3896  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/60/013,529  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LUTHER, BARBARA J.  
; REGISTRATION NUMBER: 33954  
; REFERENCE/DOCKET NUMBER: PD-0126P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 855-0555  
; TELEFAX: (415) 852-0195  
; INFORMATION FOR SEQ ID NO: 1415:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 262 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: 862319  
US-60-013-529-1415

Query Match 71.8%; Score 196; DB 66; Length 262;

Best Local Similarity 97.5%; Pred. No. 9.5e-43;

Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGAGCAGCTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCTTCCCAATGCCCCAG 60  
Db 60 ATGAGCAGCTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCTTCCCAATGCCCCAG 119  
QY 61 CTACCTCTGATACCTTCCGAGTCCGAGTGGCAGCAAAATTCGCAACCTGCTG 120  
Db 120 CTACCTCTGATACCTTCCGAGTGGCAGCAAAATTCGCAACCTGCTG 179  
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTCTCGGATGTAGTTCTCAGGTTCT 180  
Db 180 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTCTCGGATGTAGTTCTCAGGTTCT 239  
QY 181 GGCAGGGCTGCAGAAAGGCT 201  
Db 240 GGCAGGGCTGCAGAAAGGCT 260

## RESULT 13

US-08-951-197-5067  
; Sequence 5067, Application US/08951197  
; GENERAL INFORMATION:  
; APPLICANT: Gooding, Douglas H.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Akerblom, Ingrid E.  
; APPLICANT: Deleane, Angelo M.  
; APPLICANT: Naughton, Rebecca E.  
; APPLICANT: Klingler, Tod M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM HUMAN  
; NUMBER OF SEQUENCES: 5970  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA

COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/951,197  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,249  
FILING DATE: OCTOBER 1, 1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/826,438  
FILING DATE: MARCH 20, 1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/822,285  
FILING DATE: MARCH 20, 1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: CERRONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PD-0143-1 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 5067:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 250 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
IMMEDIATE SOURCE:  
CLONE: 1997229  
US-08-951-197-5067

Query Match 70.7%; Score 193; DB 14; Length 250;  
Best Local Similarity 100.0%; Pred. No. 6.2e-42;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGAGCGCTTCCCAATGCCCAAG 60  
Db 58 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGAGCGCTTCCCAATGCCCAAG 117  
Qy 61 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 120  
Db 118 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 177  
Qy 121 GGGTTGGCTCTGGTTCGGTTCGGAGCGGCGGAGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 180  
Db 178 GGGTTGGCTCTGGTTCGGTTCGGAGCGGCGGAGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 237  
Qy 181 GGCAGGGCTGCAG 193  
Db 238 GGCAGGGCTGCAG 250

RESULT 14  
US-09-539-806-30498  
Sequence 30498, Application US/09539806  
GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE

FILE REFERENCE: PD-1027 CIP  
CURRENT APPLICATION NUMBER: US/09/539,806  
CURRENT FILING DATE: 2000-03-30  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 48372  
SOFTWARE: PERL Program  
SEQ ID NO 10498  
LENGTH: 250  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: hu00555791  
US-09-539-806-30498

Query Match 70.7%; Score 193; DB 23; Length 250;  
Best Local Similarity 100.0%; Pred. No. 6.2e-42;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGAGCGCTTCCCAATGCCCAAG 60  
Db 58 ATGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGAGCGCTTCCCAATGCCCAAG 117  
Qy 61 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 120  
Db 118 CTACCTCTGATACCTTGGAGTCCGGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 177  
Qy 121 GGGTTGGCTCTGGTTCGGTTCGGAGCGGCGGAGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 180  
Db 178 GGGTTGGCTCTGGTTCGGTTCGGAGCGGCGGAGTCCGAGATGGCGGCAAAATTCGCAACCTGCTG 237  
Qy 181 GGCAGGGCTGCAG 193  
Db 238 GGCAGGGCTGCAG 250

RESULT 15  
US-09-539-806B-30498  
Sequence 30498, Application US/09539806B  
GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE  
FILE REFERENCE: PD-1027 CIP  
CURRENT APPLICATION NUMBER: US/09/539,806B  
CURRENT FILING DATE: 2000-03-30  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 48372  
SOFTWARE: PERL Program

; SEQ ID NO 30498  
; LENGTH: 250  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: hu00555791  
US-09-539-806B-30498

Query Match 70.7%; Score 193; DB 23; Length 250;  
Best Local Similarity 100.0%; Pred. No. 6.2e-42;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGGCTTCCCAATGCCCCAG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
58 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGGCTTCCCAATGCCCCAG 117  
Qy 61 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
118 CTACCTCTGATACCTTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 177  
Qy 121 GGGTTGGCTCTGGGTGCGTTGGAGGGCGGAGTCTCGGATGTAGTGTCTCAGGTTCT 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
178 GGGTTGGCTCTGGGTGCGTTGGAGGGCGGAGTCTCGGATGTAGTGTCTCAGGTTCT 237  
Qy 181 GGCAGGCTGCAG 193  
Db ||||||||||||  
238 GGCAGGCTGCAG 250

Search completed: January 31, 2005, 20:58:10  
Job time : 2461 secs

***This Page Blank (uspio)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 21:02:13 ; Search time 68 Seconds

(without alignments)  
1285.692 Million cell updates/sec

Title: US-10-057-813-13\_COPY\_253\_375

Perfect score: 123

Sequence: 1 accaagctacgtttcttca.....tgctgtcagcgaggacccc 123

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1041130

Minimum DB seq length: 0

Maximum DB seq length: 123

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23	18.7	89	4	US-09-513-999C-15513
C 2	21.4	17.4	119	4	US-09-513-999C-15214
C 3	20.6	16.7	77	4	US-09-513-999C-24985
C 4	20.4	16.6	66	4	US-09-513-999C-15004
C 5	20.4	16.6	90	4	US-08-120-324-13
C 6	20.2	16.4	106	4	US-09-513-999C-33439
C 7	20.2	16.4	116	4	US-09-513-999C-13916
C 8	19.8	16.1	112	4	US-09-513-999C-14651
C 9	19.6	15.9	92	4	US-09-513-999C-25968
C 10	19.6	15.9	103	4	US-09-513-999C-15287
C 11	19.6	15.9	103	4	US-09-513-999C-15605
C 12	19.4	15.8	78	3	US-09-463-380-7
C 13	19.4	15.8	78	4	US-09-985-357A-7
C 14	19.2	15.6	77	2	US-08-477-527A-227
C 15	19.2	15.6	77	3	US-08-481-710-227
C 16	19.2	15.6	77	5	PCT-US96-09537-227
C 17	19.2	15.6	91	4	US-09-270-767-27356
C 18	19	15.4	88	4	US-09-322-357-48
C 19	19	15.4	89	4	US-09-621-976-14962
C 20	19	15.4	112	4	US-09-702-705-1575
C 21	19	15.4	112	4	US-09-736-457-1575
C 22	19	15.4	112	4	US-09-614-124B-1575
C 23	19	15.4	112	4	US-09-671-325-1575
C 24	19	15.4	112	4	US-09-658-824-1575
C 25	18.8	15.3	96	6	5185259-5
C 26	18.6	15.1	60	3	US-08-828-712-2
C 27	18.6	15.1	60	3	US-09-063-376-2

C 28	18.6	15.1	60	4	US-09-324-782-2	Sequence 2, Appli
C 29	18.6	15.1	60	4	US-09-668-143-2	Sequence 2, Appli
C 30	18.6	15.1	90	3	US-08-952-793-176	Sequence 176, App
C 31	18.6	15.1	90	4	US-09-849-928-176	Sequence 176, App
C 32	18.6	15.1	90	5	PCT-US96-09455A-176	Sequence 176, App
C 33	18.6	15.1	95	4	US-09-513-999C-20555	Sequence 20555, A
C 34	18.6	15.1	99	4	US-09-513-999C-32775	Sequence 32775, A
C 35	18.6	15.1	100	3	US-09-298-886-20	Sequence 20, Appl
C 36	18.6	15.1	100	4	US-09-999-672-20	Sequence 20, Appl
C 37	18.4	15.0	55	4	US-08-956-171B-2677	Sequence 2677, Ap
C 38	18.4	15.0	55	4	US-08-781-986A-2677	Sequence 2677, Ap
C 39	18.4	15.0	96	1	US-08-190-802A-20	Sequence 20, Appl
C 40	18.4	15.0	96	3	US-08-477-346-20	Sequence 20, Appl
C 41	18.4	15.0	96	3	US-08-473-089-20	Sequence 20, Appl
C 42	18.4	15.0	96	4	US-08-487-072A-20	Sequence 20, Appl
C 43	18.4	15.0	107	4	US-09-513-999C-35019	Sequence 35019, A
C 44	18.4	15.0	111	4	US-09-513-999C-12508	Sequence 12508, A
C 45	18.4	15.0	122	4	US-09-513-999C-27085	Sequence 27085, A

ALIGNMENTS

RESULT 1  
US-09-513-999C-15513/c  
; Sequence 15513, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 15513  
; LENGTH: 89  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-513-999C-15513

Query Match	18.7%	Score 23;	DB 4;	Length 89;
Best Local Similarity	60.3%	Pred. No. 1.8e+02;		
Matches	38;	Conservative	0;	Mismatches 25; Indels 0; Gaps 0;
Qy	60	GCTAGACCCCTCACAGTCGCCGCGCATGTGCTGCAGTGTGGTCTGCTCAGCCGGA	119	
Db	70	GCTGTCGCTCCGCTGCGCGCAGCCGCGCGCTACTGCTGCTGCTCTCCGAT	11	
Qy	120	CCC	122	
Db	10	GCC	8	

RESULT 2  
US-09-513-999C-15214/c  
; Sequence 15214, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487





```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-120-324-13

Query Match      16.6%; Score 20.4; DB 4; Length 90;
Best Local Similarity 61.1%; Pred. No. 1.2e+03;
Matches 33; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 67 CCCCTCAGTGC GCCCCATGTGCTTCAGTGTGGTGCTCTCAGCGGGAC 120
Db 1 CTCGCCAGGGGGCGCGAGCTGCGAGCGCTGAGCGAGCTGCAAGGGCGGCC 54

RESULT 6
US-09-513-999C-33439
; Sequence 33439, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 33439
; LENGTH: 106
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-33439

Query Match      16.4%; Score 20.2; DB 4; Length 106;
Best Local Similarity 63.3%; Pred. No. 1.4e+03;
Matches 31; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 72 CACAGTGGCGCGCATGCTGCTCAGTGTGGTGCTGCTCAGCGGGAC 120
Db 46 CCCGGCGGGGCGCTGCTGAGCGGAGTGTGGTGTCTCTGGAC 94

RESULT 7
US-09-513-999C-13916
; Sequence 13916, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13916
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-13916

Query Match      16.4%; Score 20.2; DB 4; Length 116;
Best Local Similarity 54.8%; Pred. No. 1.5e+03;
Matches 40; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 49 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCGGCATGTGCTCAGTGTGGGTGCTG 108

```

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7  
OTHER INFORMATION: n=a, g, c or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 8  
OTHER INFORMATION: n=a, g, c or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 9  
OTHER INFORMATION: v=a or c or g  
US-09-513-999C-25968

Query Match 15.9%; Score 19.6; DB 4; Length 92;  
Best Local Similarity 60.4%; Pred. No. 2.2e+03;  
Matches 29; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 75 AGTGGCGCGCCATGCTGCTGAGTGTGGTGTCTGCTCAGCGGGACCC 122  
DB 1 ATTGCHNNVTGATGTCCTTGTGTGTGATCTTGGCAGGTCCCAACCC 48

RESULT 10  
US-09-513-999C-15287/c  
Sequence 15287, Application US/09513999C  
Patent No. 6783961  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duclert, A.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
FILE REFERENCE: 59 US2.REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 15287  
LENGTH: 103  
TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2  
OTHER INFORMATION: k=g or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3  
OTHER INFORMATION: k=g or t  
US-09-513-999C-15287

Query Match 15.9%; Score 19.6; DB 4; Length 103;  
Best Local Similarity 58.8%; Pred. No. 2.2e+03;  
Matches 34; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 66 CCCCTTCACAGTGGCGCCATGCTGCTGAGTGTGGTGTCTCAGCGGGACCC 123  
DB 73 CCGCTTGCCTTGGCTCCCAAGTGTGAGATTGAGCTCTGCCAGCCGCCACCC 16

RESULT 11  
US-09-513-999C-15605/c  
Sequence 15605, Application US/09513999C  
Patent No. 6783961  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duclert, A.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
FILE REFERENCE: 59 US2.REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 15605  
LENGTH: 103  
TYPE: DNA  
ORGANISM: Homo sapiens

FILE REFERENCE: 59 US2.REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 15605  
LENGTH: 103  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-513-999C-15605

Query Match 15.9%; Score 19.6; DB 4; Length 103;  
Best Local Similarity 58.6%; Pred. No. 2.2e+03;  
Matches 34; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 66 CCCCTTCACAGTGGCGCCATGCTGCTGAGTGTGGTGTCTCAGCGGGACCC 123  
DB 73 CCGCTTGCCTTGGCTCCCAAGTGTGAGATTGAGCTCTGCCAGCCGCCACCC 16

RESULT 12  
US-09-463-380-7/c  
Sequence 7, Application US/09463380  
Patent No. 6391633  
GENERAL INFORMATION:  
APPLICANT: Anne STERN; Michael BRANDT; Konrad HONOLD; Johannes AUER; Hans KOLL  
TITLE OF INVENTION: Preparation of erythropoietin by endogenous gene activation  
FILE REFERENCE: HUBR 1151 PFE/MAS  
CURRENT APPLICATION NUMBER: US/09/463,380  
CURRENT FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: PCT/EP98/04590  
PRIOR FILING DATE: 1998-07-22  
PRIOR APPLICATION NUMBER: US 09/113,692  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: DE 19753681.1  
PRIOR FILING DATE: 1997-12-03  
PRIOR APPLICATION NUMBER: EP 97112640  
PRIOR FILING DATE: 1997-07-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Wordperfect 6/7/8  
SEQ ID NO 7  
LENGTH: 78  
TYPE: DNA  
ORGANISM: Artificial sequence  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 49..60  
OTHER INFORMATION: Nucleotide sequence of the primer EX3 (Met-Gly-Ala-His).  
US-09-463-380-7

Query Match 15.8%; Score 19.4; DB 3; Length 78;  
Best Local Similarity 64.4%; Pred. No. 2.4e+03;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 70 CTCACAGTGGCGCCATGCTGCTGAGTGTGGTGTCTCAGC 114  
DB 66 CTCACAGTGGCGCCATGCTGCTGAGTGTGGTGTCTCAGC 22

RESULT 13  
US-09-985-357A-7/c  
Sequence 7, Application US/09985357A  
Patent No. 6544748  
GENERAL INFORMATION:  
APPLICANT: Anne STERN; Michael BRANDT; Konrad HONOLD; Johannes AUER; Hans KOLL  
TITLE OF INVENTION: Preparation of erythropoietin by endogenous gene activation  
FILE REFERENCE: HUBR 1151.1 CON PFE/MAS  
CURRENT APPLICATION NUMBER: US/09/985,357A  
CURRENT FILING DATE: 2001-11-02  
PRIOR APPLICATION NUMBER: US 09/463,380  
PRIOR FILING DATE: 2000-01-21

;; PRIOR APPLICATION NUMBER: PCT/EP98/04590  
;; PRIOR FILING DATE: 1998-07-22  
;; PRIOR APPLICATION NUMBER: US 09/113,692  
;; PRIOR FILING DATE: 1998-07-10  
;; PRIOR APPLICATION NUMBER: DE 19753681.1  
;; PRIOR FILING DATE: 1997-12-03  
;; PRIOR APPLICATION NUMBER: EP 97112640  
;; PRIOR FILING DATE: 1997-07-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: Wordperfect  
;; SEQ ID NO 7  
;; LENGTH: 78  
;; TYPE: DNA  
;; ORGANISM: Artificial sequence  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 49..60  
;; OTHER INFORMATION: Nucleotide sequence of the primer EX3 (Met-Gly-Ala-His).  
US-09-395-357A-7

Query Match 15.8%; Score 19.4; DB 4; Length 78;  
Best Local Similarity 64.4%; Pred. No. 2.4e+03;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
QY 70 CTCACAGTGGCGCCCATGTGCTGCAGTGTGGTGTCTCAGC 114  
Db 66 CTCACGTCGGCGCCCATGTGCTGCAGTGTGGTGTCTCAGC 22

RESULT 14  
US-08-477-527A-227/c  
; Sequence 227, Application US/08477527A  
; Patent No. 5972599  
; GENERAL INFORMATION:  
; APPLICANT: DIANE TASSET  
; APPLICANT: NIKOS PAGRATIS  
; APPLICANT: SUMEDHA JAYASENA  
; APPLICANT: LARRY GOLD  
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS  
; TITLE OF INVENTION: OF CYTOKINES  
; NUMBER OF SEQUENCES: 258  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Swanson and Bratschun, L.L.C.  
; STREET: 8400 East Prentice Avenue, Suite #200  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,527A  
; FILING DATE: 7-JUNE-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/714,131  
; FILING DATE: 10-JUNE-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/931,473  
; FILING DATE: 17-AUGUST-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/964,624  
; FILING DATE: 21-OCTOBER-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/117,991  
; FILING DATE: 8-SEPTEMBER-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/536,428  
; FILING DATE: 11-JUNE-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Diane H. McClearn

;; REGISTRATION NUMBER: 33,960  
;; REFERENCE/DOCKET NUMBER: NEX41-1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (303) 793-3333  
;; TELEFAX: (303) 793-3433  
;; INFORMATION FOR SEQ ID NO: 227:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 77 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA  
US-08-477-527A-227  
Query Match 15.6%; Score 19.2; DB 2; Length 77;  
Best Local Similarity 54.2%; Pred. No. 2.8e+03;  
Matches 39; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 32 GCTGGTCCAGCCTCACAGGCTAGACCCCTCACAGTGGCGCCCATGTGC 91  
Db 77 GCGGATCCGGCCTCATGTGGAACAGTGCAGCTGAGCGGACTGCGCATCGACTTGAGC 18  
QY 92 CTCAGTGTGGG 103  
Db 17 GTTATTCTGAG 6  
RESULT 15  
US-08-481-710-227/c  
; Sequence 227, Application US/08481710  
; Patent No. 6028186  
; GENERAL INFORMATION:  
; APPLICANT: DIANE TASSET  
; APPLICANT: NIKOS PAGRATIS  
; APPLICANT: SUMEDHA JAYASENA  
; APPLICANT: LARRY GOLD  
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS  
; TITLE OF INVENTION: OF CYTOKINES  
; NUMBER OF SEQUENCES: 258  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Swanson and Bratschun, L.L.C.  
; STREET: 8400 East Prentice Avenue, Suite #200  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/481,710  
; FILING DATE: 7-JUNE-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/714,131  
; FILING DATE: 10-JUNE-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/931,473  
; FILING DATE: 17-AUGUST-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/964,624  
; FILING DATE: 21-OCTOBER-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/117,991  
; FILING DATE: 8-SEPTEMBER-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/536,428  
; FILING DATE: 11-JUNE-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Diane H. McClearn  
; REGISTRATION NUMBER: 33,960

;/ REFERENCE/DOCKET NUMBER: NEX41-2  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (303) 793-3333  
;/ TELEFAX: (303) 793-3433  
;/ INFORMATION FOR SEQ ID NO: 227:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 77 base pairs  
;/ TYPE: nucleic acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: DNA  
US-08-481-710-227

Query Match 15.6%; Score 19.2; DB 3; Length 77;  
Best Local Similarity 54.2%; Pred. No. 2.8e+03;  
Matches 39; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 32 GCTGGTCCAGCCTCACCTGACACAGGGCTAGACCCCTCACAGTGGCGCCGCGCATGTGC 91  
Db 77 GCCGGATCCGGGCTCATGTGGAACAGTGCAGCTGAGGGGACTGCGCATCGACTTGAGC 18  
QY 92 CTGCAGTGTGGG 103  
Db 17 GTTATTCTGAG 6

Search completed: January 31, 2005, 22:17:25  
Job time : 69 secs